



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : H04Q	A2	(11) International Publication Number: WO 00/51359 (43) International Publication Date: 31 August 2000 (31.08.00)
(21) International Application Number: PCT/US00/04934 (22) International Filing Date: 25 February 2000 (25.02.00) (30) Priority Data: 09/258,487 26 February 1999 (26.02.99) US (71) Applicant: MATCHLOGIC, INC. [US/US]; 10333 Church Ranch Boulevard, Westminster, CO 80021 (US). (72) Inventor: MCELHINEY, James, D.; 114 Fourth Avenue, Ottawa, Ontario K1S 2L4 (CA). (74) Agents: YOUNG, James, R. et al.; Chrisman, Bynum & Johnson, P.C., 1900 Fifteenth Street, Boulder, CO 80302 (US).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: METHOD FOR DETERMINING GEOGRAPHIC LOCATION OF USERS CONNECTED TO OR USING A NETWORK		
(57) Abstract A method for determining the geographic location of users connected to a network and for determining the accuracy or confidence level of such geographic location determination includes the steps of collecting data about or regarding users connected to or using a network, determining the geographic location of such users using multiple techniques, determining or computing a confidence level for each geographic location technique used for each user, receiving a request for geographic location information for one of the users, selecting the geographic location technique to be used for the requested user, and sending geographic location information for the specified user to the requester of the information.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Method For Determining Geographic Location of Users Connected to or Using a Network

Technical Field

This invention relates to a method for determining the geographic location of a user of a computer or other communications network and, more specifically, to a method for determining a geographic location of a user of a computer or other communications network and the accuracy of such determination.

Background Art

During recent years there have been rapid advancements in computers and computer networking. In particular, the worldwide network of computers commonly referred to as the Internet has seen explosive growth. The Internet comprises a vast network of smaller wide area and local area computer networks connected together so as to allow the sharing of resources and to facilitate data communication between computers and users. The rapid growth of the Internet is due, in large part, to the introduction and widespread use of graphical user interfaces called browsers which allow users easy access to network servers and computers connected to the Internet and, more particularly, the World Wide Web.

The World Wide Web forms a subset of the Internet and includes a collection of servers, client devices, computers, and other devices. Each server may contain documents formatted as web pages or hypertext documents that are accessible and viewable with a web compliant browser, such as the Netscape Navigator™ or Communicator™ browsers or the Mosaic™ browser. Each hypertext document or web page may contain references to graphic files or banners that are to be displayed in conjunction with the hypertext document or web page. The files and banners may or may not be stored at the same location as the hypertext document or web page.

A hypertext document often contains hypertext links to other hypertext documents such that the other hypertext documents can be accessed from the first hypertext document by activating the hypertext links. The servers connected to the World Wide Web utilize the Hypertext Transfer Protocol (HTTP) which is widely known protocol which allows users to use browsers to access web pages and the banners or files associated with web pages. The files, banners, hypertext documents, or web pages may contain text, graphics, images, sound, video, etc. and are generally written in a standard page or hypertext document description language known as the Hypertext Markup Language (HTML). The HTML format allows a web page developer to specify the location and presentation of the graphic, textual, sound, etc. on the screen displayed to the user accessing the web page. In addition, the HTML format allows a web page to contain links, such as the hypertext links described above, to other web pages or servers on the Internet. Simply by selecting a link, a user can be transferred to the new web page, which may be located very different geographically or topologically from the original web page.

When using a conventional browser, a user can select which web page or hypertext document the user wishes to have displayed on the user's computer or terminal by specifying the web page's Universal or Uniform Resource Locator (URL) address. Each server has a unique URL address and, in fact, so does each web page and each file needed to display the web page. For example, the URL address for the U.S. Patent and Trademark Office is currently <http://www.uspto.gov>. When a user types in this URL address into a browser, the user's terminal establishes a connection with the U.S. Patent and Trademark Office and the initial web page for the U.S. Patent and Trademark Office is transmitted from the server storing this web page (which may or may not be actually located at the U.S. Patent and Trademark Office) to the user's terminal and displayed on the user's terminal. The web page may include a number of graphic images or elements, often referred to as banners, which are to be displayed on the user's terminal in conjunction with the web page. Each of the graphic images is typically stored as a separate file on the server and has its own URL address. When the web page is initially transmitted from the server to the user's terminal, the browser receives the URL addresses for the graphic images and then requests that they be transmitted from the server on which they are stored to the user's terminal for display on the user's terminal in conjunction with the web page. The server(s) on which the graphic images are stored may or may not be the same server on which the original web page is stored. More specifically, since the URL's addresses for the included graphic images are all processed separately using the HTML protocols, it is possible and, in fact, common, for these graphic images to be stored on separate and even widely distributed computers or hosts, all of which are accessible to the user's terminal via a computer network. For purposes of the present invention, the term "banner" is meant to be construed very broadly and includes any information displayed in conjunction with a web page wherein the information is not part of the same file as the web page. That is, a banner includes anything that is displayed or used in conjunction with a web page, but which can exist separately from the web page or which can be used in conjunction with many web pages. Banners can include graphics, textual information, video, audio, animation, and links to other computer sites, web sites, web pages, or banners.

The growth of easy access to the World Wide Web and the ability to create visually pleasing web pages have helped increase the amount of advertising and other promotional materials created for use and display with web pages. For example, a car manufacturer may have a web page describing the company and the cars and car parts that the company manufactures and sells. Part of the web page may include advertising information or banners such as, for example, images of current car models sold by the manufacturer or the types and numbers of cars the manufacturer has in stock. The car manufacturer may also contract with the owners or operators of other web pages to have the car manufacturer's advertisement banners displayed when users access these other web pages. Similarly,

an advertising agency may contract with various web sites to have the advertisement banners of the agency's clients displayed when users access the web pages stored on the web sites. For example, an advertising agency or ad-network firm may contract with a web site containing general information about cars to have advertising information or banners included on the web pages displayed to a user
5 accessing the web site. The advertising banners may contain graphics, text, etc. about car models or car parts manufactured by one of the advertising agency's clients. Furthermore, the advertisement banners may not be stored on the same server or computer or web site on which the web page is stored. Rather, all or a significant portion of the advertisement banners created by an advertising agency may reside on one or more information or ad servers. Typically, an advertising agency will pay a fixed
10 amount of money for a fixed number of displays of its advertisement banners on a single web page or group of web pages. Therefore, advertising agencies are understandably very interested sending or displaying advertising banners to users that are as geographically relevant or targeted to the user as possible. For example, if the user is found or determined to be located in Colorado, U.S.A., the advertiser may wish to send an advertising banner that relates or is directed to Colorado, U.S.A. If the
15 advertiser can determine that the user is located in Westminster, Colorado, U.S.A., as opposed to Denver, Colorado, U.S.A., or Boulder, Colorado, U.S.A., the advertiser may wish to send a different advertising banner that is targeted or related to the city of Westminster, Colorado, U.S.A., instead of the broader first advertising banner that is directed to Colorado, U.S.A.

Unfortunately, the state of the art is such that accurate determination of the geographic
20 location of users connected to or using a computer network is not available. Therefore, despite the well-developed state of the art in the sending and displaying of information, banners, and advertisements in conjunction with web pages, documents, or other information, there is still a need for a method of determining the geographic location of a user or users connected to or using a computer network and of determining the accuracy of such geographic location determination for the user or
25 users. The method will preferably be able to recognize and deal with anomalous or incorrect geographic location information that might be associated with one or more of the users and be able to recognize and deal with a user connected to a computer network using network addresses, such as internet protocol (IP) addresses that float or change over time or between subsequent connections of the user to the computer network.

30 **Disclosure of Invention**

Accordingly, a general object of the present invention is to provide a method for determining geographic location of users connected to or using a computer or other communications network.

Another general object of the present invention is to provide a method for determining the accuracy of each method used to determine the geographical location of users connected to or using a computer or other communications network.

Yet another general object of the present invention is to provide a method for determining geographic locations of users connected to a computer or other communications network that reduces the impact of user having anomalous geographic location information associated with them.

A further general object of the present invention is to provide a method for determining the geographic location of a user connected to or using a computer other communications network at varying levels of geographic specificity or hierarchy.

Still another general object of the present invention is to provide a method for selecting among various techniques or reason lineages for determining the geographic location of a user connected to or using a computer or other communications network.

Another general object of the present invention is to provide a method for determining geographic locations of users connected to a computer or other communications network using fixed or floating network addresses.

Additional objects, advantages, and novel features of the invention shall be set forth in part in the description that follows, and in part will become apparent to those skilled in the art upon examination of the following or may be learned by the practice of the invention. The objects and the advantages may be realized and attained by means of the instrumentalities and in combinations particularly pointed out in the appended claims.

To achieve the foregoing and other objects and in accordance with the purposes of the present invention, as embodied and broadly described herein, the method of the present invention includes collecting data about or regarding users connected to or using a network and the network addresses used by the users when they are connected to or logged on to the network determining the geographic location of such users using multiple techniques, determining or computing a confidence level for each geographic location technique used, receiving a request for geographic location information for one of the users, selecting the geographic location technique to be used for the requested user, and sending geographic location information for the specified user to the requester of the information.

Also to achieve the foregoing and other objects and in accordance with the purposes of the present invention, as embodied and broadly described herein, the method of the present invention includes collecting information regarding geographic location of a first set of one or more users of the network, each individual user in said first set having at least one network address that can be associated to said individual user's geographic location information; determining a pool of network addresses, said pool of network addresses containing at least one network address associated with a particular user in

said first set of users; and establishing at least a portion of said particular user's geographic location information as geographic location information for all users of network addresses in said pool of network addresses.

Also to achieve the foregoing and other objects and in accordance with the purposes of the present invention, as embodied and broadly described herein, the method of the present invention includes collecting geographic location information for a user associated with at least one network address; determining a first network address for which a cookie or other specific identifier is or has been associated and a second network address, distinct from and higher than said first network address for which said cookie or other specific identifier also is or has been associated; associating a counter to at least one network address between said first network address and said second network address; and incrementing each of said at least one counters.

Also to achieve the foregoing and other objects and in accordance with the purposes of the present invention, as embodied and broadly described herein, the method of the present invention includes selecting a range of network addresses; selecting a first network address within said range of network addresses; determining if a specific identifier has been associated with at least one network address in said range of network addresses that is higher than said first network address and if said specific identifier has been associated with at least one network address in said range of network addresses that is lower than said first network address; and repeating said process for a second network address within said range of network addresses.

Also to achieve the foregoing and other objects and in accordance with the purposes of the present invention, as embodied and broadly described herein, the method of the present invention includes determining with a first technique a first possible geographic location of the specific user; determining with a second technique a second possible geographic location of the specific user; determining whether said first technique or whether said second technique provides a more accurate approximation of the specific user's actual geographic location; and calibrating said first technique and said second technique.

Also to achieve the foregoing and other objects and in accordance with the purposes of the present invention, as embodied and broadly described herein, the method of the present invention includes gathering geographic location information for at least one user who has been connected to the network; determining a first set of network addresses that have been used or associated with said at least one user; determining a first set of cookies or other specific identifiers that have been used or associated with any network addresses in said first set of network addresses; determining a second set of network addresses that have been used or associated with any of said cookies or other specific identifiers in said first set of cookies or other specific identifiers; determining a second set of cookies

that have been used or associated with any of said network addresses in said second set of network addresses; and repeatedly determining sets of cookies or other specific identifiers and network addresses until such time as said sets of network addresses stabilizes for consecutive determinations.

5 Also to achieve the foregoing and other objects and in accordance with the purposes of the present invention, as embodied and broadly described herein, the method of the present invention includes gathering geographic location information for at least one user who has been connected to the network; determining a lower network address and a higher network address that have been used or associated with said at least one user; determining a first set of cookies or other specific identifies that have been used or associated with any network addresses in a range between said lower network
10 addresses and higher network address; determining a new lower network address and a new higher network address that have been used or associated with any of said one or more specific identifiers in said first set of one or more specific identifiers; determining a second set of one or more specific identifiers that have been used or associated with any network addresses in a range between said new lower network address and set new higher network address; and repeatedly determining sets of specific
15 identifiers and lower and higher network addresses until such time as said lower network address stabilizes and said higher network address stabilizes.

Brief Description of the Drawings

The accompanying drawings, which are incorporated in and form a part of the specification, illustrate the preferred embodiments of the present invention, and together with the descriptions serve to explain the principles of the invention.

5 Figure 1 illustrates the method of the present invention;

 Figure 2 illustrates a computer network over which the method of the present invention illustrated in Figure 1 can be implemented;

 Figure 3 illustrates the strict float pool technique of performing the step of determining geographic locations of users in the method of Figure 1;

10 Figure 4 illustrates the range float pool technique of performing the step of determining geographic locations of users in the method of Figure 1;

 Figure 5 illustrates the boundary float pool of performing the step of determining geographic locations of users in the method of Figure 1;

15 Figure 6 is an representative histogram generated during use of the boundary float pool of Figure 5;

 Figure 7 is another representative histogram generated during use of the boundary float pool of Figure 5;

 Figure 8 is a further representative histogram generated during use of the boundary float pool of Figure 5;

20 Figure 9 illustrates the proxy server float pool technique of performing the step of determining geographic locations of users in the method of Figure 1;

 Figure 10 is a representative histogram generated during the use of the boundary float pool technique of Figure 5 wherein the histogram does not accurately show the true boundaries of the float pool;

25 Figure 11 is another illustration of the histogram of Figure 10, shown with a higher degree of granularity than the histogram of Figure 10;

 Figure 12 is another representative histogram generated during use of the boundary float pool technique of Figure 5;

30 Figure 13 is another representative histogram generated during use of the boundary float pool technique of Figure 5;

 Figure 14 is another representative histogram generated during use of the boundary float pool technique of Figure 5; and

 Figure 15 is an illustration of a calibration process that can be used in conjunction with the method of Figure 1.

Brief Description of the Drawings

A method 30 in accordance with the principles of the present invention is illustrated in Figure 1 and includes step 32 of collecting data and information regarding one or more users of a computer network, step 34 of determining the geographic location of one or more users connected to the computer network using one or more techniques or sub-techniques, also referred to as reason lineages, based on the information collected during the collection step 32, step 36 of determining or establishing the accuracy or confidence level of each technique or sub-technique used to determine the geographic location of users, step 38 of receiving a request for geographic location information for a specific user, step 40 of selecting which technique or sub-technique among the various techniques or sub-techniques used during the step 34 to determine the geographic location of users connected to the computer network will be used to answer the request received during the step 38 for geographic location information for a specific user connected to the computer network, and step 42 of supplying geographic location information about the user specified during the step 38 to the requester based on the geographic location determination technique or sub-technique selected during the step 40. The geographic location information about the specified user sent during the step 40 can be used for a variety of things, including use in the selection of which advertisement or banner to serve the specified user. Each of these steps will be discussed in more detail below. In general, the method 30 attempts to determine the geographic locations of at least some users connected to a computer network, such as the Internet or World Wide Web, and then determines which other groups of users connected to the computer network can be assumed to be in the same geographic locations as these users.

A significant feature of the method 30 of the present invention is that it allows the use of various geographic location determination techniques and sub-techniques, also referred to as reason lineages, for users connected to a computer network during the step 34 and then preferably selects from the various techniques and sub-techniques during the step 40 the technique or sub-technique for a specific user that is the most precise or accurate based upon the determination during the step 36 of the confidence level or accuracy of each geographic location determination technique or sub-techniques used. Therefore, geographic location data collected about users connected to computer network can come from a variety of sources or be generated by a variety of techniques and sub-techniques and the method 30 does not require any one specific method of geographic location data collection for the users.

Another significant feature of the method 30 of the present invention is that it allows the geographic location of users connected to or using a computer network to be determined at various levels of geographic specificity. For example, the geographic location of each user can be viewed at several levels of geographic hierarchy or specificity, including the planet, continent, region country, state/province, major city or metropolitan area, town, area code or telephone exchange, ZIP or postal

code, etc. for the user. Each level of geographic specificity may have a different confidence level associated with it. For example, the method 30 may determine with an eighty percent (80%) confidence level that a particular user is in the United States, but may only determine with a forty percent (40%) confidence level that the user is in the state of Colorado. As another example, the method 30 may
5 determine with a sixty percent (60%) confidence level that a particular user is in Europe, but may only be able to determine with a twenty percent (20%) confidence level that the user is in Spain.

A significant advantage provided by the method 30 is that advertisers, banner servers, etc. can send advertisements to users connected to a computer network that are geographically targeted to the user. For example, if an advertiser knows with an eighty percent (80%) confidence level that a
10 particular user connected to a computer network is in Denver, Colorado, U.S.A., the advertiser may send or serve an advertisement or banner to the user that is specifically directed to such geographic location. If, on the other hand, the advertiser has only a ten percent (10%) confidence level that the user is in Denver, Colorado, U.S.A., but the advertiser has a fifty percent (50%) confidence level that the user is in Colorado, U.S.A., the advertiser might send a more broadly directed advertisement to the
15 user. That is, the advertiser can send or serve the user an advertisement or banner that is relevant to Colorado, U.S.A., but that is not narrowly tailored to Denver, Colorado, U.S.A.

With the ability to determine at various confidence levels the continent, country, state/province, county, time zone, ZIP code, city or metropolitan area, etc. of users connected to a computer network, advertisers can target advertisements sent to users based on their desired criteria. Thus, an advertiser
20 can send one of several different kinds of advertisements or banners based on the level of confidence determined for each level of geographic location specificity. For example, an advertiser might only send advertisements that have a minimum confidence level for the level of geographic specificity. Therefore, the advertiser might only send a user an advertisement directed to a particular city or metropolitan area if the advertiser has at least a minimum confidence level that the user is located in or
25 near the particular city or metropolitan area, else the advertiser might send the user a broader advertisement directed to a particular state/province or country if there is a higher confidence level at the broader levels of geographic specificity. Each of these advantages and features will be discussed in more detail below.

Due to the rapid growth of computer networks such as the Internet and the World Wide Web,
30 users may be connected to a computer network 50 from geographically diverse locations, as best illustrated in Figure 2. For example, users may be connected to the computer network 50 via computers, terminals, or other client devices 52, 54 and may be located in Denver, Colorado, U.S.A., and Westminster, Colorado, U.S.A., respectively or computers, terminals, or other client devices 56, 58 and may be located in Ottawa, Ontario, Canada, and Montreal, Quebec, Canada, respectively. In

addition, users may be connected to the computer network 50 via client devices or computers 60, 62 which may be located in the metropolitan area of New York City, New York, U.S.A., and the metropolitan area of New York City, New Jersey, U.S.A. respectively. Other users may be located in Sydney, Australia, Paris, France, Cairo, Egypt, and Al-Minyā, Egypt, and may access the computer network 50 via terminals or computers 63, 64, 66, 67, respectively. In addition to users connected to the computer network 50, servers, such as the servers 68, 70, 72, and the proxy server 74 and other devices or servers may also be connected to the computer network 50.

The computer network 50 may constitute or include wide area networks, local area networks, intranets, the Internet, the World Wide Web, etc. and is not limited by the type of network or network topology. The computer network 50 illustrated in Figure 1 is only meant to be generally representative of computer networks for purposes of elaboration and explanation of the present invention and other client devices, servers, networks, etc. may be connected to the computer network 50 without departing from the scope of the present invention. The computer network 50 is also intended to be representative of, and include, the Internet, the World Wide Web, privately or publicly owned or operated networks such as, for example, Tymnet, Telenet, America On-Line, Prodigy, Compuserve, Information America, and the Microsoft Network, and other local or wide area computer networks. The computer network 50 can also include or be representative of corporate or other private intranets, which are privately owned networks using Internet protocols.

For purposes of elaboration and explanation, but not limitation, of the present invention, the conventions and protocols of the Internet, the World Wide Web, and browsers therefore, will be used as examples, in particular, the concept of a Uniform Resource Locator (URL), the Hypertext Transfer Protocol (HTTP), the Hypertext Markup Language (HTML), Internet Protocol (IP) addresses, and the Transmission Control Protocol/Internet Protocol (TCP/IP). In addition, for purposes of explanation, but not limitation, of the method 30 of the present invention, the computer network 50 will be considered to be the Internet, World Wide Web, or other computer network using similar protocols as the Internet and the World Wide Web. It should be noted, however, that the concepts underlying the present invention can be used for computer networks using other or different types of architectures, client/server models, conventions, protocols, and network addressing schemes. For more details on these protocols, the reader is directed to: Kevin Washburn and Jim Evans, TCP/IP running a successful network, 2nd Ed. (1996), published by Addison-Wesley, Douglas E. Comer, Internetworking with TCP/IP, 3rd Ed. (1995), published by Prentice Hall, John December and Mark Ginsberg, HTML 3.2 and CGI Unleashed Professional Reference Edition (1996), published by Sams.net Publishing, and Jerry Honeycutt *et al.*, Using HTML 3.2, 3rd Ed (1997), published by Que Corporation. Other information about the HTTP, HTML, TCP/IP and other network protocols can also be found in U.S.

Patent No. 5,617,540 issued to *Civanlar et al.*, U.S. Patent No. 5,572,643 issued to *Judson*, and U.S. Patent No. 5,442,771 issued to *Filepp et al.*

It should also be noted that the disclosed method also works for all types of operating systems running on the computers, terminals, computer sites, information servers, and other devices connected to the computer network 30. Such operating systems can include, for example, Microsoft Corporation's DOS™, WINDOWS 3.x™, WINDOWS NT™, WINDOWS 95™, or WINDOWS 98™ software, IBM's OS/2™ software, Apple's System 7™ software, Sun Corporation's Solaris™ software, or the AIX or UNIX operating system software platforms.

Now referring to Figure 1, computers or terminals can be connected to the computer network 50 in a variety of ways. For example, computers, terminals, or other client devices can be connected directly to the computer network 50 or may be attached via a dial-up line or network access service provider. Other client devices, computers, or terminals 76, 78 may be connected to the computer via network proxy or local servers, such as the proxy server 74. Proxy servers allow multiple computers, terminals, or other client devices to be connected to a computer network at a single point. For example, a large corporation may have all its client devices and servers connected via a local area computer network. The local area computer network can be connected to a caching proxy server which is, in turn, connected to the computer network 50. In the computer network 50 illustrated in Figure 1, the client devices 76, 78 access the computer network 50 through the proxy server 74. Similarly, the client devices 80, 82, 84 access the computer network 50 through the proxy server 86. Using proxy servers allows multiple client devices access to a computer network while limiting the number of physical connections between the client devices and the computer network.

Preferably, the computer network 50 is based on the Internet Protocol (IP) which designates a unique address for each device connected to the computer network 50 and defines a scheme for giving each such device a unique address. In a computer network based on the Internet Protocol, the IP address for a particular device is not based on the type of device or computer network, how the particular device operates, or how what the device is connected to. In such an implementation of the Internet Protocol (IP), each computer or web site and other host devices, end systems, networks, or network router devices connected to the computer network 50 has a unique Internet Protocol (IP) address that is thirty-two bits in length and is generally written as four decimal numbers in the range zero (0) through 255, separated by periods. For example, an IP address could be 128.10.2.30, which in its full thirty-two bit format is 10000000.00001010.00000010.00011110. Currently over four billion IP addresses are available. Later versions of the Internet Protocol (IP), such as IPv6, will enhance the current Internet Protocol scheme and allow a larger number IP addresses. Providing every computer or other device on a computer network with a unique IP address allows any host computer to communicate

with any other host computer and allows for terminals or computers connected to the computer network 50 to be identified.

While some devices connected to the computer network 50 will have fixed Internet Protocol (IP) addresses that do not vary over time, a significant fraction of users and devices connected to the computer network 50 will not have fixed IP addresses. Instead, these users will have IP addresses that are dynamically assigned to them at the time the users log on, connect to, or establish a session with the computer network 50, typically via an internet service provider (ISP) using a protocol such as the Dynamic Host Configuration Protocol (DHCP). DHCP is a protocol for automatic TCP/IP configuration that provides static and dynamic IP address allocation and management.

When a user logs on to the computer network 50 via an internet service provider, thereby commencing a connection or browser session, the user is assigned an IP address by the internet service provider that is fixed for the duration of the user's connection session. When the user terminates the session by disconnecting from the computer network 50, the IP address assigned to the user is freed up and becomes available for reallocation or reassignment by the internet service provider to another user that commences or initiates a session. Thus, IP addresses are dynamically assigned by the internet service provider to users connecting to the computer network 50 and allowed to "float" between users connecting to the computer network 50, thereby allowing a finite number of IP addresses to be used by a potentially greater or even infinite number of users over time.

Internet protocol (IP) addresses are generally assigned by Internet Assigned Number Authority (IANA) which has ultimate control over assignment and allocation of IP addresses and the Internet Network Information Center (INTERNIC) which can also assign or allocate IP addresses. Not all IP addresses are allocated directly by an authority. For example, the United States authority allocated a large block of IP addresses to Apnic (Asia Pacific Network Information Center), which in turn has allocated sub-blocks of the IP addresses to major internet service providers (ISPs). The internet service providers may, in turn, allocate, sub-sub-blocks of IP addresses to smaller firms. As a result of such allocation of IP addresses, no central list or database exists as to allocated IP addresses.

Internet service providers (ISPs) are typically assigned a fixed block or pool of IP addresses from which dynamic allocation or assignment of IP addresses to users logging onto or connecting to the computer network 50 can be made. Thus, each user's IP address, taken over a series of days, months, etc., will dynamically float around the bounds of the IP address pool allocated to the specific internet service provider through which the user connects to the computer network 50. Dynamic allocation of IP addresses may also happen with private companies. For example, an employee for a company that dials or logs in to the company multiple times from home or other remote locations might be allocated or assigned a different IP address each time by the company's computer system, each of the user's

temporarily allocated or assigned IP address being within the range of IP addresses previously allocated or assigned to the company. The use, allocation, and operation of the Internet Protocol and IP addresses are well known to people of ordinary skill in this art and need not be explained in any further detail for purposes of explanation of the present invention.

5 The method 30 of the present invention preferably takes advantage of information that is known about users connected to the computer network 50 and that can be related to the IP addresses that the users obtain, even if only temporarily, when the users connect to the computer network 50. For example, a cookie can be used to relate a particular user to the floating or dynamically assigned IP addresses assigned to the user when the user connects to the computer network 50. Therefore, the
10 cookie becomes associated with the IP addresses used by the user. When a user uses a web browser at a computer, terminal, or other connection or client device, such as the computer 52 in Figure 2, to access or establish a session with a server, such as the web server 70 in Figure 2, via a computer network using TCP/IP and HTML protocols, the user's web browser typically sends an information or serve page request to the web server. The web server will answer the request by sending or serving the
15 desired information to the user's computer for display on the computer by the web browser. The web server will often also generate a cookie or set-cookie command and send it along with the requested information or web page to the user's browser such that a cookie is then be stored on the user's computer. The user's web browser will then send the cookie back to the web server when sending subsequent requests for information or pages to the web server.

20 Many different kinds of information can be stored in the cookie such as a user identification number, a client account number associated with the user, the time and date, and the expiration time for the cookie (*i.e.*, the length of time the cookie will remain valid). As a user connects and reconnects to a computer network over a period of time and sends requests to the same web server or to one in a group of web servers, the cookie stored on the user's computer may be sent by the browser on the user's
25 computer to the web server during or as part of each request, thereby associating the cookie with the IP addresses temporarily or permanently assigned to the user. Thus, the cookies can be used to detect and monitor the IP addresses used by the user during the period of time. The use and operation of cookies and user identification numbers embedded in cookies are well known to people of ordinary skill in this art and need not be explained in any further detail for purposes of explanation of the present invention.

30 Referring now again to Figure 1, the method 30 of the present invention will now be discussed in more detail. A significant feature of the method 30 is that users' Internet Protocol (IP) addresses are used to determine the geographic location of the users. As previously discussed above, certain assumptions can be made about the allocation and use of IP addresses that are related to the geographic location of users connected to the computer network 50, as will be discussed in more detail below.

During the collection step 32, information and data regarding users and Internet Protocol (IP) addresses associated with or assigned to the users are collected from a wide variety of sources. For purposes of the present invention, users will be preferably categorized into three different classes. Class one users are users for which the names and addresses of the users are known and which have been tagged to a unique identification number, such as a cookie identification number. Class one data or information includes the name, address, etc. information for the user. Class one user information or records can be gathered from a number of sources but are typically expensive to gather. For example, contests or games conducted via the computer network 50 might require that a user enter the user's name and address to enter the contest or to play the game. Thus, the user's IP address can be directly associated with a geographic location. Class two users are users for which the names and addresses are not known, but some other transaction history exists such that an anonymous profile exists. For example, the transaction history information might include a cookie or other distinct or specific identifier associated with the IP address of a class two user. Class three users are users for which only a minimum amount of information is known. When a user generates a page, banner, or information request signal and sends it across the computer network 50 to a web server or other device, the user's IP address, host name, browser type, operating system, and referring pages's Universal or Uniform Resource Locator (URL) address may be sent with or as part of the request signal as part of the HTTP request header. While the user's host name is not generally included, each transaction with or request from the user generally includes the user's IP address that can be associated with any geographic location information then available or previously determined.

During the collection step 32, it is preferable that cookies or other kinds or types of distinct or specific identifiers for users be monitored by web servers and other devices connected to the computer network 50 such that information regarding IP addresses assigned to users can be stored and relationships or associations between known cookies and IP addresses can be developed. The result of the collection step 32 will be a database or set of information regarding some or all possible Internet Protocol (IP) addresses and the geographic, cookie or other distinct identifiers, and other information associated or used with each possible IP address. Class one, two, and three records or information may also be collected and associated, if possible, with IP addresses. The information collected for different IP addresses may vary widely and have different levels of accuracy, as will be discussed in more detail below.

After the collection step 32 is completed for a given period of time T, the collected user, cookie, and IP address data is used during the step 34 to determine the geographic location of users connected to the computer network 50, as will now be discussed in more detail. Preferably, the time T is between one and six months and is optimally approximately three months or ninety days. Many different

techniques and sub-techniques can be used to determine the geographic location of individual users connected to the computer network 50 during the step 34 and the method 30 of the present invention is not limited to any particular geographic determination technique or sub-technique. In fact, during the step 34, multiple techniques and/or sub-techniques are preferably used to determine the geographic locations of the users.

Many different kinds of techniques or reason lineages are possible for determining the geographic information of users connected to the computer network 50. Each technique may be preferably broken down into sub-techniques, as will be discussed in more detail below. An example geographic determination technique, domain name registration information is publicly available that can be associated with particular IP addresses. When a user sends a request signal across the computer network 50 to a server or other device connected to the computer network 50, the user's IP address is generally available or included with or as part of the request signal. Via a process of reverse DNS (domain name system) look up, approximately eighty-five percent (85%) of the time, the user's host name can be determined from the user's IP address. The user's domain name will generally form part of the user's host name. Therefore, from the user's host name, the domain name can be determined and the registration authority that issued the domain name can be found. There is approximately one domain name registration authority per country, many of which will make domain name registrant information available. Thus, the registrant of the domain name can often be determined along with the registrant's telephone number and/or address. The telephone number and/or address of the registrant can be used to create assumptions about the user. Note that the accuracy of these assumptions for different users may vary widely. For example, domain names assigned to large companies are not necessarily indicative of where users connected to the computer network 50 via the host computer associated with the domain name are geographically located. In addition, large internet service providers, such as America On-Line, will provide access to the computer network 50 to users in a large geographic area, so the domain registration information for the large internet service provider will not always be indicative of the locations of users connecting to the computer network 50 via the large internet service provider. If the user's host name is not included in the request signal sent by the user, a *gethostbyaddr* routine can be performed which uses the user's IP address to search the computer network 50 to find the name of the corresponding computer and the computer's host name. Once the host name has been determined, the technique continues as described above.

Another technique to determine the geographic location of a user during the step 34 is based on netblock registrations of IP addresses. As previously discussed above, IP addresses are generally assigned or allocated in blocks of 256 consecutive IP addresses or multiples thereof. Blocks of IP addresses often, but not always, correspond to internet service providers. Information regarding the

top level IP address allocations is publicly available. Therefore, when a user sends a request signal, the owner of the IP address block in which the user's IP address falls can be determined. The telephone number and/or address of the owner of the IP block can be used as described above in regard to the domain name registration technique to determine a probable geographic location of the user. As with
5 the domain name registration technique, the accuracy of the assumptions made for users based on ownership of the IP addresses may vary widely, particularly since large internet service providers may use different subsets of their allocated IP addresses in different geographic regions of the world or of a particular country.

A third technique for determining geographic information for users connected to the computer
10 network 50 uses the traceroute (sometimes called tracert) feature of the HTTP protocol. The traceroute technique is a well known feature for determining the likely path through a computer network, such as the Internet or the World Wide Web, between two points or devices connected to the computer network. By using the *gethostbyaddr* routine on each intermediate result, the names of the intermediate routers between the two points can be determined. Many routers for the Internet include geographic location
15 information encoded into their names. The geographic location information from the routers can be used to determine geographic location information for the IP addresses. Problems associated with this technique include the fact that it can be slow to complete and is not one hundred percent (100%) theoretically sound. In addition, half to two-thirds of routers will not have immediately identifiable geographic location information encoded in their names and, even if they do contain such encoded
20 geographic location information, the information may be difficult to parse given the millions of different abbreviations and formats in use for encoding geographic location information.

The three techniques previously described above for providing geographic location information for a user connected to the computer network 50 during the step 34 are well known to people having ordinary skill in the art and need not be described any further for purposes of elaboration of the method
25 30 of the present invention. However, new techniques can also be used to determine geographic location information of users during the step 34. One such technique, referred to as the "float pool" technique, and its variations will now be discussed in more detail.

A float pool can be thought of as a set of more than one contiguous IP address which is used as a common pool of IP addresses for a set of more than one cookie. That is, a float pool consists of IP
30 addresses, usually consecutive, each of which has seen more than one of the same cookies or which have been associated with more than one of the same cookies. Testing and other empirical evidence shows that, statistically, the users of a float pool of IP addresses are highly likely to be geographically located near each other. One reason that this is generally true is that routing is done by splitting and resplitting ranges of IP addresses. Therefore, it is very convenient to have all of the IP addresses for a

physical location or geographic area to be contiguous. The equipment which dynamically allocates IP addresses is usually located, at least topologically, just before the "last mile," *i.e.*, next to the bank of modems, which people usually call into using local-access, not long distance, telephone calls.

Therefore, the pool of dynamically assigned IP addresses does not usually server a geographic area larger than the size of a local telephone calling area. If a record or database of IP addresses and cookies associated with those IP addresses is maintained for a period of time, say ninety days, certain assumptions can be made about users based on the IP addresses used by the users. The collection or database of IP addresses and associated cookies can be created over the time period by monitoring or analyzing requests for pages, information, banners, etc. sent by users or client devices to web servers or other devices connected to the computer network 50.

The basic theory for the float pool technique as used for geographic location determination or targeting during the step 34 is based on the following observations. First, most internet service providers (ISP) are small or large. Small internet service providers typically have relatively localized geographic coverage, usually one or a few cities or other limited area. The IP addresses used by such small internet service providers are often allocated from a single pool of consecutive IP addresses. Large internet service providers usually have many IP addresses, but for purposes of optimizing the routing of information or data packets within computer networks, and of allowing local devices to manage pools of IP addresses at dial-in or dial-up centers, the large internet service providers typically break up large blocks of IP addresses into many smaller blocks, each block of which is used primarily within a finite geographic area.

In the case of either a small or large internet service provider, the geographic location of all users within a pool of IP addresses is usually highly correlated. Therefore, given a sample of class one records for a given pool of IP addresses, a profile or database can be built regarding the users of all of the IP addresses within the pool of IP addresses.

The first variation of the float pool technique is the strict float pool technique 98, as best illustrated in Figure 3. In the strict float pool technique 98, a user or IP address for which class one information is known or associated is chosen during step 100. Then, all IP addresses used by the user during a certain time period are determined during step 102 from the relationship between cookies stored on the user's computer and IP addresses for the user at which the cookie was seen or associated. Next, during step 104, cookies are collected which used or were associated during the given time period with any of the IP addresses in the set of IP addresses determined during the step 102. During step 106, IP addresses which ever had or were ever associated with any cookies from the collected set of cookies from step 104 are determined for the given time period. A determination is made during step 107 if the set of IP addresses is stable. That is, a determination is made whether the set of IP addresses

determined during the step 106 is different from the previously determined set of IP addresses (which on the first pass is the set of IP addresses determined during the step 102). If the determination made during the step 107 is negative, *i.e.*, consecutively determined sets of IP addresses are different, then the set of IP addresses is not stable and the list of cookies for the new set of IP addresses is determined during step 108 and the process is repeated. If the determination made during the step 107 is affirmative, *i.e.*, consecutively determined sets of IP addresses are identical, then the set of IP addresses is assumed to be stable and forms a float pool of IP addresses. Statistically, any user of any IP address within the float pool of IP addresses is likely to be in the same general geographic location as the original user(s) whose class one information was used to start the process. Alternatively, if other class one users have used IP addresses that fall within the float pool of IP addresses, the level of geographic hierarchy (*i.e.*, planet, continent, country, region, state/province, time zone, etc.) which they have in common may be assigned to the entire float pool of IP addresses. Since float pools are usually contiguous ranges of IP addresses, all IP addresses within the minimum and maximum IP address are assumed to be in the float pool, even if some of the IP addresses did not occur during the iterative process of steps 102, 104, 106, 108. The process is then preferably repeated for the next user or IP address for which class one information is known by selecting a new class one user during step 109. The IP address for which class one information exists and chosen during the step 109 is checked to see if it is already within a known float pool of IP addresses. If not, the process is repeated beginning at the step 102. If the class one user or IP address chosen during the step 109 is already within a known float pool of IP addresses, the technique 98 preferably checks again to see if float pools for all IP addresses with associated class one records have been determined, as previously discussed above. The technique is preferably continued until float pools of IP addresses for all users or IP addresses for whom class one records exist are determined or computed.

The strict float pool technique 98 has some limitations. For example, the strict float pool technique 98 is memory intensive and the sets and databases of cookies and IP addresses can become quite large. Moreover, "cookie leakage" can occur which generates spurious results. "Cookie leakage" occurs when a cookie is sent by widely varying IP addresses and can occur for many reasons. For example, cookies sent by web servers to specific users, though intended to be unique, may not be unique in all circumstances. For example, in some cases, many users may get assigned the identical cookie, thereby creating an association between the cookie and a larger set of IP addresses. Another reason for cookie leakage, possibly the most common reason, is that a user simply uses the same computer to connect to a computer network from a different internet service provider, or moves to a new city. Another reason for cookie leakage is that an internet service provider may change the way it which the IP addresses within its purview or control are managed or allocated during a sampling period.

Cookie leakage may also occur when a user signs up for a nationwide or international dial-up account with an internet service provider, like Netcom. In this case, the user will be seen in widely varying IP addresses over time. Cookie leakage causes float pools of IP addresses to merge into other pools and makes it possible for all users to be assigned to a single giant float pool by accident.

5 A second variation of the float pool technique is the range float pool technique 110 and is best illustrated in Figure 4. The range float pool technique 110 is very similar to the strict float pool technique 98 and relies in a database of IP addresses and associated cookies built up over a given period of time. In the range float pool technique 110, a user or IP address for which class one information is known is chosen during step 112. Then, the minimum and maximum IP addresses used
10 by the user during a given period of time are determined during step 114. Next, cookies which used or which were associated with any of the IP addresses for a given period of time in the range between the minimum and maximum IP addresses determined during the step 114 are collected during step 116. Then the minimum and maximum IP addresses which ever had any cookies from the collected set of cookies from step 116 is determined during step 118. A determination is made during step 119 if the
15 minimum and maximum IP addresses are stable. That is, a determination is made during the step 119 as to whether the minimum and maximum IP addresses determined during the step 118 are different from the previously determined minimum and maximum IP addresses (which on the first pass are the minimum and maximum IP addresses determined during the step 114). If the determination made during the step 119 is negative, *i.e.*, consecutively determined minimum and maximum IP addresses are
20 different, then the set of IP addresses is not stable and the list of cookies for the new set of IP addresses is determined during step 120 and the process is repeated. If the determination made during the step 119 is affirmative, *i.e.*, consecutively determined minimum and maximum IP addresses are identical, then the minimum and maximum IP addresses are stable and the range of IP addresses within the minimum and maximum IP addresses forms a float pool of IP addresses. Any user of any IP address
25 within the float pool of IP addresses can be posited or shown statistically to be in the same geographic location as the original user whose class one information was used to start the process. Alternatively, if other class one users have used IP addresses that fall within the float pool, the level of geographic hierarchy (*i.e.*, planet, continent, region, country, state/province, major metropolitan area, town, time zone, etc.) which they have in common can be assigned to the entire float pool of IP addresses. The
30 process is then repeated for the next user about whom class one information is known by selecting a new class one user during step 122 and so on until float pools for all class one users are determined or computed.

Like the strict float pool technique 98 previously discussed above, the range float pool technique 110 also has some limitations. For example, the range float pool technique 110 is also

memory intensive, although not as memory intensive as the strict float pool method. In addition, the range float pool technique 110 can be more susceptible to cookie leakage than is the strict float pool technique 98.

A third variation of the float pool technique is the restricted range float pool technique 130 which is a modification of the range float pool technique 110. Empirical evidence and observation indicates that very few float pools of IP addresses are over 4,096 IP addresses in size and, therefore, 4,096 is preferably chosen during use of the restricted range float pool technique 130. Therefore, in the restricted range float pool technique 130, all cookies that have been seen at IP address ranges exceeding 4,096 IP addresses are eliminated from consideration during the steps 116 and 120 of the range float pool technique 110. Thus, cookies which would tend to create large float pools are eliminated from consideration. The 4,096 limit on the allowed range of IP addresses is variable and can be set to other desired limits. However, limitations still exist in the restricted range float pool technique 130. For example, the restricted range float pool technique 130 can still be memory and time extensive.

A fourth and preferred technique for determining geographic location of users during the step 34 is the boundary float pool technique 140, which is best illustrated in Figure 5. The boundary float pool technique 140 is based on the assumption that a float pool of IP addresses can be considered as a set of IP addresses, each IP address of which has the property that a finite set of known cookies have appeared or been associated with IP addresses in the set both below and above it. The known cookie may be different for each of the IP addresses in the set. This assumption ignores the actual minimum and maximum IP addresses forming the boundary of the float pool of IP addresses. In the boundary float pool technique 140, a counter is created for each possible IP address. During step 140 each of the IP address counters are given an identical initial starting value, such as zero (0).

During step 144, a determination is made for each cookie or portion of cookie regarding the minimum and maximum IP address at which the cookie or portion of cookie was found or associated. During step 146, each IP address counter corresponding to IP addresses between the minimum and maximum IP address determined for each cookie during step 144 is incremented by one. The resulting pattern formed by the histogram of IP address counters can be analyzed during step 148 to determine the IP addresses corresponding to float pools of IP addresses and the geographic location to be associated with the float pools can be determined during step 149. For example, histograms or graphs 150, 151, 152 illustrated in Figure 6, respectively, show exemplary results of steps 142, 144, 146 for different ranges of IP addresses. Histogram 150 illustrates the counters for IP addresses 195.232.2.0 to 195.232.31.224, histogram 151 illustrates the counters for IP addresses 195.232.33.0 to 195.232.62.224, and histogram 152 illustrates the counters for IP addresses 195.232.65.0 to 195.232.79.224. The histograms 150, 151, 152 illustrate the number of distinct cookies that has

appeared above and below each of the IP addresses in the histograms 150, 151, 152. The histograms 150, 151, 152 represent three large float pools of IP addresses while the histograms 153, 154, 155 in Figure 6 represent three small float pools of IP addresses. Additional representative histograms of float pools are shown in Figure 7. Histogram 156 represents a large float pool for the range of IP addresses between 195.77.80.0 to 195.77.95.224. Histograms 157, 158, 159 illustrate smaller float pools of IP addresses. Histogram 161 illustrated in Figure 8 is a good example of a float pool of IP addresses which was probably moved from IP address 207.16.5.96 (CF100560 in hexadecimal format) to 207.16.8.32 (CF100820 in hexadecimal format). Since the intervening IP addresses are not all the same value, they are probably still in use.

In the simplest case, histograms resulting from steps 142, 144, 146 will be zero between float pools of IP addresses and will have a relatively constant, non-zero high value within the span of a float pool of IP addresses. In reality, however, vagaries in the way that the Internet is used will create many unusual cases and irregular histograms. For example, histograms 150, 151, 152 shown in Figure 6 are rounded instead of flat because relatively few people have appeared at IP addresses at the edges of the float pools of IP addresses represented by the histograms 150, 151, 152 and relatively more people have appeared at IP addresses at both above and below the middle of the float pools represented by the histograms 150, 151, 152. In general, a change of four or more between a given IP address counter and the counters within a window of twenty IP address counters on either side of the given IP address counter is a good indicator of the boundary of a float pool of IP addresses for use during step 148.

If desired, an optional step may be included between the steps 144 and 146 of the boundary float pool technique 140 that removes all cookies from consideration that have appeared at a range of IP addresses whose maximum IP address is no more than 4,096 IP addresses above its minimum IP address. As previously discussed above, empirical evidence and observation indicates that very few float pools are over 4,096 IP addresses in size. Therefore, in the boundary float pool technique 140, all cookies that have been seen at IP address ranges exceeding 4,096 IP addresses are preferably eliminated from consideration after the step 144 and before the step 146. The 4,096 limit on the allowed range of IP addresses is variable and can be set to other desired limits.

As previously discussed above, once float pools of IP addresses have been determined during step 148, class one data for each of the float pools of IP addresses is preferably analyzed during step 149 to determine the geographic location to be associated with users of the IP addresses with the float pool of IP addresses. One way to assign the geographic location for IP addresses in a given float pool of IP addresses is to look at all class one information associated with each IP address in the float pool of IP addresses and find the lowest level of geographic hierarchy common to the entire float pool of IP

addresses. For example, referring once again to histogram 150 in Figure 6, assume that class one records exist for five of the IP addresses within the float pool of IP addresses represented by the histogram 150, as identified by cookies. The lowest level of common geographic location (*i.e.*, continent, country, area code, time zone, major metropolitan area, state/province, etc.) between each of the five IP addresses can be assigned to each of the other IP addresses in the float pool. Should no class one information exist for any of the IP addresses in the float pool of IP addresses, then the float pool of IP addresses is not considered further. If desired, an optional and intermediate step 169 can be performed after the step 148 and before the step 149 which extends the range of contiguous IP addresses into a larger multi-address pool record in order to better recognize float pools of IP addresses having ranges of contiguous IP addresses which do not fall on the sampled IP address boundaries. For example, suppose a range of contiguous IP addresses with one hundred distinct cookies, thereby passing the threshold tests at step 148, starts at IP address 198.205.100.0 and ends at IP address 198.205.100.255. The boundary float pool technique 140 may still not accurately detect the start and end IP addresses of the float pool because: (1) the boundary float pool technique 140 may be run for every two, thirty-two, etc. IP addresses, not every IP address, thereby introducing small, but conservative, errors in the estimation of the exact limiting IP address of each float pool; and (2) the boundary float pool technique 140 counts only those cookies that were seen both above and below a given IP address; therefore, the example pool from IP address 198.205.100.0 to IP address 198.205.100.255, sampled at every thirty-two IP addresses, might be found to extend only from IP address 198.205.100.32 to IP address 198.205.100.224. For this reason, each pool of IP addresses is extended downwards to the next lower sample boundary, and upwards to the IP address that is one less than the next higher sample boundary, to arrive at the more likely correct IP address boundary locations. Note that, unlike the float pool techniques 98, 110, 130 previously described above, the boundary float pool technique 140 does not use class one data or information as its starting point. Thus, the boundary float pool technique 140 produces lists of all possible float pools of IP addresses at the same time, while only those float pools of IP addresses having at least one IP address with an associated one class one record are assigned a geographic location. An advantage of this is that one may thereby obtain or create an estimate or confidence level of the coverage or completeness of the geographic database thus constructed. The assignment of confidence levels to float pools will be discussed in more detail below.

As previously described above, many users may connect to or access the computer network 50 via proxy servers, such as the proxy server 74 illustrated in Figure 2. In such a case, each of these users will have the same IP address which constitutes a float pool of a single IP address. Although such single-IP-address float pools are properly detected by the strict float pool technique 98, the

boundary float pool technique 140 is usually unable to detect these pools of IP addresses, since the boundary float pool technique 140 does not process cookies seen at a single IP address. In order to detect such proxy servers, a proxy server float pool technique 160 is preferably used, as best illustrated in Figure 9.

5 In the proxy server float pool technique 160, the number of distinct cookies seen for each IP address over a fixed period of time is counted during step 162. All IP addresses have their number of distinct cookies falling below a threshold are then discarded during step 164. Preferably, the threshold used during the step 162 is high enough to eliminate users with multiple cookies (such as users who have reinstalled their browsers, switched computers, or deleted their cookie file from time to time) but
10 low enough to catch all proxy servers with meaningfully large numbers of users. For example, empirical testing has shown that a threshold number of eight or more cookies at a specific IP address works well for the step 164. After the particular IP addresses are discarded during step 164, all IP addresses falling in a multi-IP address float pool are discarded during step 166 and the remaining single IP address float pools are designated as proxy servers during step 168. The multi-IP address float
15 pools can be determined using the boundary float pool technique 140 previously described above.

 After the steps 162, 164, 166 are completed, the remaining IP addresses can be considered to represent single IP-address proxy servers. Should class one data exist for any of the IP addresses designated as proxy servers, all users using the proxy server's IP address can be considered to be in the geographic location provided by the class one data.

20 When float pools of IP addresses are created using either the restricted range float pool technique 130 or the boundary float pool technique 140 previously discussed above, the boundaries for the float pools of IP addresses as determined by the methods will not always correspond with actual float pool boundaries. For example, now referring to Figure 10, an exemplary histogram or float pool of IP addresses 180 is shown that might have been generated by the boundary float pool technique 140.
25 The edges or boundaries of the float pool 180 of IP addresses are shown at 182 which approximately corresponds to IP address 199.3.72.32 (C7034820 in hexadecimal format) and at 184 which approximately corresponds to IP address 199.3.75.224 (C7034BE0 in hexadecimal format). However, in this example, the points 182, 184 do not provide the true edges of the float pool 180 of IP addresses. This is because the granularity of the histogram 180 is such that it only shows IP addresses which are a
30 multiple of thirty-two. Histogram 185, as illustrated in Figure 11, represents the same float pool as that represented by the histogram 180, but has edge boundaries 186, 188 which correspond to IP address 199.3.72.10 (C703480A in hexadecimal format) and IP address 199.3.75.248 (C7034BF8 in hexadecimal format), respectively, due to showing IP addresses which are a multiple of two. Thus, the histogram 185 is a more accurate representation of a float pool of IP addresses than is the histogram

180. Note that, in the first case, the optional step 169 previously described above would extend the pool of IP addresses from IP address 199.3.72.0 to IP address 199.3.75.255. In the second case, with a resolution of two IP addresses, optional step 169 would extend the pool of IP addresses from IP address 199.3.72.8 to IP address 199.3.75.250. With or without the optional step 169, the resolution of two IP addresses gives a more accurate result for the boundary of the float pool, but makes little difference to the overall accuracy of the geographic targeting method 30 or a database formed from the use of the method 30 in the absence of another float pool which is immediately adjacent in IP address.

Further examples of histograms 190, 192, 194 that could be generated using the boundary float pool technique 140 are illustrated in Figures 12, 13, 14, respectively. If thirty-two IP addresses resolution is used, Figure 12 represents a float pool of IP addresses extending from IP address 199.67.189.32 (C743BD20 in hexadecimal format) to IP address 199.67.196.96 (C743C460 in hexadecimal format). Using two IP address resolution (not illustrated in Figure 12), this same float pool of IP addresses will be found more accurately to extend from IP address 199.67.189.6 (C743BD06 in hexadecimal format) to IP address 199.67.196.100 (C743C464 in hexadecimal format).

Other situations when histograms may not reflect the true boundaries of a float pool are when, for example, cookies did not appear with IP addresses at the highest and lowest possible IP address in the float pool. In addition, such a situation might occur when the time period over when samples are taken is not long enough, or the float pool in question is not heavily used.

As previously discussed above, IP addresses are generally allocated in blocks of 256. Therefore, for IP addresses which are above or below a multi-address float pool of IP addresses or a single IP address float pool, thereby not falling within a float pool of IP addresses, but which are in the same aligned block of 256 IP addresses will be considered to be "float puddles." An aligned block of IP addresses can be defined as a set of IP addresses that start with the same three numbers when expressed in "dotted" notation, *e.g.*, 205.189.80.xxx which includes all IP addresses from 205.189.80.0 through 205.189.80.255 inclusively. Algebraically, an aligned block of IP addresses means all thirty-two bit IP addresses "A" of the form $(256 \times N) \leq A \leq (256 \times (N+1) - 1)$ where N is any integer in the range $0 \leq N \leq 16,777,215$.

Float puddles of IP addresses will preferably be given the same geographic location as the float pools of IP addresses to which they are adjacent, but at a slightly lower confidence level, as will be discussed in more detail below. As a further example of a float puddle of IP addresses, assume that an internet service provider owns the IP address block 205.189.78.0 through 205.189.78.255, a block of 256 IP addresses, and also assume that a float pool of IP addresses has been determined to exist between the IP address range 205.189.78.30 through 205.189.78.200, then two float puddles of IP

addresses can be defined which are located on either side of the float pool of IP addresses and have the IP address ranges 205.189.78.0 through 205.189.78.29 and 205.189.78.201 through 205.189.78.255.

While the float pools techniques 98, 110, 130, 140, 160 described above work well to generate float pools of IP addresses, there may be instances when blocks of IP addresses do not float. Currently, approximately fifteen percent (15%) of transactions on the World Wide Web are currently believed to be generated by users using IP addresses which are fixed and approximately eighty-five percent (85%) are currently believed to be generated by users using IP addresses that float. In a fixed IP address, the same IP address represents the same machine or device day after day and the machine or device does not typically use more than one IP address. Empirical testing has shown, however, that the usage of IP address blocks, due to IP routing conventions with masks for routing decision points, means that the IP addresses within a generally fixed IP block are often physically very close to one another. Therefore, a new type, called a fixed pool, of pool of IP addresses can be defined as follows. If class one data exists for one or more IP addresses within an aligned block of 256 IP addresses, but no float pool of IP addresses is determined to exist within the aligned block of 256 IP addresses, then the aligned block of 256 IP addresses is treated as a float pool of IP addresses, but is called a fixed pool of IP addresses so as to be able to distinguish it from the float pools generated by the techniques 98, 110, 130, 140, 160 previously described above. The lowest geographic hierarchy common to all IP addresses within the fixed pool for which class one data exists can be used, if desired, for the all of the IP addresses within the fixed pool of IP addresses. Each such fixed float pool of IP addresses so generated is added to the list of float pools of IP addresses and float puddles of IP addresses generated by the techniques previously described above and a record is preferably kept of the technique used to generate each such range of IP addresses so that the accuracy of the technique can be determined and calibrated, as will be discussed in more detail below.

After geographic locations for users are determined during the step 34 using the techniques described above, as well as other possible techniques, a confidence or accuracy level or value is preferably assigned during step 36 to each of the possible techniques used during step 34. The confidence level or value preferably provides a quantified assessment of the accuracy of the technique. For example, a confidence level of eighty percent (80%) for the strict float pool technique 98 for a given user is an indication that the user is eighty percent (80%) likely to be geographically located where the results of the strict float pool technique 98 have predicted the user to be. More particularly, consider the case that one-hundred class one persons with known IP addresses have been seen at an IP address determined to be within a particular pool (whether it be a boundary-type float pool, proxy-type float pool, or fixed pool). Also assume further that the single most common town among those one-hundred users is the town of Westminster, a suburb of Denver, Colorado, USA, and that forty of the

one-hundred live in that town and further that a total of sixty persons live within the confines of Greater Denver (forty of whom live in Westminster), and a total of ninety persons live within Colorado (sixty of whom are in Denver, and forty of whom are in Westminster). In this case, the geographic targeting information and associated confidences for this pool of IP addresses would be as follows:

5	State or Province	Colorado	Confidence = 90%
	Major Metropolitan Area	Greater Denver	Confidence = 60%
	Town	Westminster	Confidence = 40%

By performing a large number of such confidence calculations, it is possible to come up with an average accuracy for the technique being used, without reference to any single particular pool of IP addresses. For example, after considering many such float pools of IP addresses, it might be determined that the average accuracy for the float pool technique when used with over one-hundred class one records as follows:

	State or Province	Any	Confidence=85%
	Major Metropolitan Area	Any	Confidence=65%
15	Town	Any	Confidence=35%

This same technique of calibration of accuracy can be applied to any geographic targeting technique for training purposes, provided that a collection of class one records (users with known addresses) can be found which are able to be targeted using the geographic targeting technique in question. For example, given a number of class one records for a given domain registration, one is able to calculate the likelihood that a user is in the same town, metropolitan area, state, etc., as the address contained within the registration data for that internet domain. This might be further broken down into targeting methods such as domain registration for small companies, domain registration for large companies, domain registration locations for internet service providing companies, etc., or any number of classifications of targeting method as might be available to a practitioner expert in the field. Thus, it is possible to determine, for each type of targeting or geographic location determining technique, for each level of statistical support, and for each level of the geographic hierarchy, an approximate expected accuracy level.

The step 36 may be done prior to or simultaneously with step 34, in whole or in part, but for purposes of clarity in elaboration of the method 30, the step 36 will be considered to be performed after completion of the step 34.

In general, assigning a confidence level or value during the step 36 to each geographic location determination technique or reason lineage used during the step 34, such confidence level being indicative of the accuracy of the technique, allows the method 30 to choose between which technique to use at a given time for determining during step 40 the geographic location of a specific user and

providing such user location information to a requester of such information during the step 42. Quite often, more than one geographic targeting or location determination technique will be available to determine the geographic location of a specific user. For example, the strict float pool technique 98 and a technique based on netblock registrations, as previously described above, may generate location information for a specific user during the step 34. Assigning a confidence level or value to each of the two techniques during the step 36 allows a choice to be made during the step 40 as to which technique provides the highest probability of yielding a correct answer.

In order to assign to determine confidence levels for each user location finding technique, the techniques used during the step 34 are preferably further broken down into sub-techniques and confidence levels are made specific to the data or sub-technique in question. For example, the strict float pool technique 98 may be broken down into several sub-techniques based on the number of class one sample points and level of geographic hierarchy that the sample points have in common. If the method 30 is assumed to have a geographic hierarchy limited to continent, country, region, state/province, major metropolitan area, town, ZIP code, and telephone exchange, a minimum of eight subsets of the strict float pool technique 98 may be found. That is, a first subset of the strict float pool technique 98 exists in the case that all class one records used to test the float pool in question have the same continent, but differ in their country; a second sub-technique of the strict float pool technique 98 exists in the case that all class one records used to test the float pool in question have the same continent and country, but differ in their regions; a third sub-technique of the strict float pool technique 98 exists in the case that all class one records used to test the float pool in question have the same continent, country, and region, but differ in their state or province, etc. Thus, for this example, the strict float pool technique 98 is broken down to eight possible further sub-techniques, each sub-technique of which is directed to a specific level of geographic hierarchy and preferably has a confidence level or value for each level of the geographic hierarchy, as will be described in more detail below. For example, a specific user whose location is determined using the third sub-technique described above tested with class one records common to the region but with differing states or provinces) might have a higher confidence at the region level than at the state or province level.

In addition to breaking down a location finding technique into sub-techniques by level of geographic hierarchy, the method 30 also preferably breaks down each technique into sub-techniques by number of class one user records available that are common at the different levels of geographic hierarchy. Continuing the example previously described above, different float pools of IP addresses determined using the strict float pool technique 98 may contain different numbers of class one records. That is, for example, one float pool may contain a single IP address associated with a single class one record, while a different float pool may contain IP addresses associated with ten class one records.

Therefore, more geographic location information is available for the second float pool than for the first float pool, presumably allowing a higher confidence level to be assigned to the second float pool than to the first float pool if all of the class one records for the second float pool contain identical geographic information. Thus, while the strict float pool technique 98 is one technique for determining a float pool, such float pool might contain IP addresses for users having class one records that are common to the continent, region, state/province, major metropolitan area, etc., levels and for each of those levels, there may be a different number of class one records. Therefore, the method 30 preferably breaks down each geographic location finding or determination technique 98, 110, 130, 140, 160 into subsets or sub-techniques by the levels of geographic hierarchy, each subset or sub-technique further broken down by the numbers of class one records available to each float pool or fixed pool of IP addresses. For the eight levels of commonality of geographic hierarchy used in the previous example (*i.e.*, continent, country, region, state/province, major metropolitan area, town, ZIP code, and telephone exchange), each of the eight subsets or sub-techniques of the strict float pool technique 98 based on geographic hierarchy is preferably further broken down into float pools including IP addresses for which only one class one record exists, float pools including IP addresses for which two to three class one records exist, float pools including IP addresses for which four to seven class one records exist, float pools including IP addresses for which eight to fifteen class one records exist, and float pools including IP addresses for which sixteen or more class one records exist. Therefore, the strict float pool technique 98 can be broken down into forty subsets including a subset or sub-technique directed to the continent of a specific user in a float pool having an IP address for which a single class one record is available in which the continent is found, a subset or sub-technique directed to the country of a specific user in a float pool having a range of IP addresses for which two or three class one records are available in which the country is found in common, a subset or sub-technique directed to the major metropolitan area of a specific user in a float pool in a range of IP addresses for which eight through fifteen class one records are available in which the major metropolitan area is found in common, etc. The other techniques 110, 130, 140, 160 are preferably broken down in to similar subsets or sub-techniques based on levels of commonality of geographic hierarchy and numbers of available class one records.

While each of the techniques is broken down so that each subset or sub-technique is directed to a level of geographic hierarchy, each subset or sub-technique preferably is also represented by a vector or array containing accuracy or confidence percentages or levels for all of the levels of geographic hierarchy. For purposes of further elaboration, each of the subsets or sub-techniques of the techniques 98, 110, 130, 140, 160 will be referred to as a reason or technique lineage. Therefore, a reason lineage refers to a particular technique or sub-technique of determining geographic location of a user or users.

Each reason or technique lineage preferably has a vector or array associated with it, with each entry in the vector or array relating to a different level of geographic hierarchy. Each of the reason lineage vectors or arrays can be assigned a confidence level or percentage at each level of geographic hierarchy which represents the accuracy percentage that the reason lineage has correctly predicted or determined the level of geographic hierarchy directly associated with it and the remaining levels of geographic hierarchy as well. For example, for a reason lineage directed to the state level of geographic hierarchy and a sample size of eight to fifteen class one records, the reason lineage vector might contain a seventy percent (70%) confidence at the state level of geographic hierarchy that the state has been correctly determined by the reason lineage, a one-hundred percent (100%) confidence that the continent, country, and region have been correctly determined by the reason lineage, since they are usually known with complete accuracy once the state is known, a thirty percent (30%) confidence level at the major metropolitan area level of geographic hierarchy that the major metropolitan area has been correctly determined by the reason lineage, and a zero percent (0%) confidence level for the ZIP code and telephone exchange levels of geographic hierarchy.

As previously described, each of the reason lineages or sub-techniques of the techniques 98, 100, 130, 140, 160 provides confidence or accuracy information for all levels of the geographic hierarchy in the reason lineage vector, but the actual confidences in the reason lineage vector will vary based on (a) the number of class one records used to determine the location of the pool in question, and (b) the variance in the geographic locations of the class one records so considered. For example, if the user is determined to be in a float pool of IP addresses in which five class one user records are available, and those five class one users were all in the same major metropolitan area for a reason lineage vector directed to a state level of geographic hierarchy, but in differing towns, the user might obtain confidences of eighty percent (80%) for the continent, country, region, state, and major metropolitan area, but confidences of zero percent (0%) for town and postal code. Such a reason lineage vector indicates that there is an eighty percent (80%) confidence that the reason lineage correctly determined the continent, country, region, state, and major metropolitan areas for the user and a zero percent (0%) confidence that the reason lineage correctly determined the town and postal code for the user. Alternatively, if the user is determined to be in a float pool of IP addresses in which one hundred class one users were all in the same major region but in differing states, (and therefore differing major metropolitan areas, towns, and postal codes), the user might obtain confidences of ninety-five percent (95%) for the continent, country, and region, but confidences of zero percent (0%) for state, major metropolitan area, town and postal code. Therefore, the reason lineage vector indicates that there is an ninety-five percent (95%) confidence that the reason lineage correctly determined the continent, country, and region for the user and a zero percent (0%) confidence that the reason lineage

correctly determined the state, major metropolitan area, town, and postal code for the user. Thus, a larger number of class one sample records for the pool of IP addresses will often increase the confidence of the levels of the geographic hierarchy which are in common for those class one records, and there will be little or no confidence for those levels of the geographic hierarchy which vary among the class one records used to calibrate the float pool of IP addresses in question.

Reason lineage vectors also preferably exist for the other geographic location determination techniques described above that are not based on pools of IP addresses. For example, the domain name registration information technique previously described above may be broken down by level of geographic hierarchy, by size of the company, and/or by whether the domain name ends in “.net” or “.org” or “.com”. Presumably a company with a small number of employees is more likely than a company with a large number of employees to have all or a significant majority of its employees in the same general geographic location. Therefore, a geographic location determination technique based on domain name registration might have different sub-techniques, each of which is specific to a particular type of domain, or type or size of organization, or method of geographic location (*e.g.*, by telephone number of registration information, or by ZIP code), and each such sub-technique has a reason lineage vector which described the geographic accuracy for all levels of the geographic hierarchy specific to all domains using that specific sub-technique. Geographic location determination technique based on netblock registrations, traceroute routines, etc., can also be broken down into sub-techniques, each such sub-technique having reason lineage vector specific to its particular sub type, but containing confidence levels or percentages for each of the levels of the geographic hierarchy. It should be noted that there is not a reason lineage for every specific float pool and every netblock and every specific domain, etc., but rather one for each sub-type of float pool, and one for sub-type of netblock, etc., such that each vector of confidences represents an average accuracy level across all specific float pools (or netblocks, etc.) of that specific sub-type of float pool (or netblock, etc.).

In addition to reason lineage vectors, the method 30 preferably includes a set or database of place lineage vectors, each place lineage vector of which relates a specific geographic location to the other geographic locations in the geographic hierarchy. Thus, for a given place at a specific level of geographic hierarchy, the place lineage vector provides percentages at the other levels of geographic hierarchy. For example, a place lineage vector for a place representing “Colorado” at the state/province level of geographic hierarchy for a user would have a confidence level of one-hundred percent (100%) for continent, country, region, and state/province. The place lineage vector for Colorado might also contain an educated guess that the major metropolitan area for the user is Denver with a confidence level of seventy percent (70%), based on the fact that most people who live in Colorado, U.S.A., also live in the Denver metropolitan area. The place lineage vector for Colorado

might also contain confidence levels of zero percent (0%) for ZIP code and telephone exchange since they vary widely throughout the state of Colorado, U.S.A. A place lineage vector for the Denver metropolitan area of Colorado, U.S.A., would contain a confidence level of one-hundred percent (100%) for the continent, country, region, state/province, and major metropolitan area, since each of these levels of geographic hierarchy are known once the major metropolitan area is known. The place lineage vector might also contain a confidence level of zero percent (0%) for the ZIP code given that most people living in the Denver metropolitan area have widely varying ZIP codes, and a confidence level of fifty percent (50%) for the town (also Denver) since only a fraction of the inhabitants of the greater Denver area live within the city limits of the incorporated City of Denver proper.

As a second example of a place lineage vector, a place lineage vector for a place representing "New York City" might have an confidence level of one-hundred percent (100%) for continent, country, major metropolitan area, and time zone, but lower confidence levels for state/province, county, and telephone exchange since they vary widely for the city of New York, U.S.A. As another example, a known telephone exchange may serve one, two, or more ZIP codes. If the telephone exchange is in an area with only one ZIP code, the place lineage vector for the telephone exchange will have a confidence of one-hundred percent (100%) for the ZIP code. If the telephone exchange is in an area serving three ZIP codes, the place lineage vector for the telephone exchange might have a confidence of thirty percent (30%) for any one of the three ZIP codes.

Thus, like reason or technique lineage vectors, each place lineage vector is initially directed to a specific type of geographic location that corresponds to one of the eight levels of geographic hierarchy for which confidence levels are calculated. The remainder of the place lineage vector provides the percent likelihood that a person is in the associated locations of other types or hierarchy levels, given that the user has been found to be in the known initial geographic location. In all cases, the confidence or accuracy of at least one level of the geographic hierarchy in a place lineage vector will be one-hundred percent (100%), at the level of the geographic hierarchy corresponding to the type of place to which the place lineage vector is associated. For example, the place lineage vector for a state is always one-hundred percent (100%) accurate to the state level of geographic hierarchy, and the place lineage vector for a town is always one-hundred percent (100%) accurate at the town level of geographic hierarchy.

Place lineage vectors can be determined by reference to atlases, maps, census data, and other demographic information. Obviously, confidence levels for geographic hierarchy for some place lineage vectors will never change. For example, the state of Colorado will almost certainly always also be in the United States of America and in North America. However, telephone exchanges and ZIP codes for the state of Colorado will change over time.

For the strict float pool method 98, range float pool method 110, restricted range float pool method 130, and boundary float pool method 140 previously described above, class one data can be used to determine the geographic location of users using IP addresses from a float pool of IP addresses, even if the class one data used is not the same for all levels of geographic hierarchy. For example, suppose that use of the boundary float pool method 140 finds the float pool of IP addresses illustrated by the histogram 161 in Figure 8. Now assume that within the float pool of IP addresses defined by the histogram 161, one hundred class one users were known to visit IP addresses in the range of IP addresses in the pool. Ideally, the geographic information for all of the one hundred users of those IP addresses in the pool. That is, the class one information would indicate that users for each of the one hundred IP addresses would be in the same continent, region, country, state/province, major metropolitan area, town, ZIP code, and telephone exchange. Generally, however, the class one information for IP addresses within a float pool of IP addresses will vary at multiple levels of geographic hierarchy. This variance may be for a variety of reasons, including imperfect data arising from the techniques used to collect the class one data. While a goal for the method 30 is high accuracy, perfect accuracy is not absolutely required. Therefore, the float pool techniques 98, 110, 130, and 140 preferably set a threshold level for each level of geographic hierarchy. Using the previous example with the one hundred IP addresses for which class one information is known, assuming the threshold level is eighty-five percent (85%), if, for each level of geographic hierarchy, eighty-five percent (85%) of the class one samples have the same result, that result is taken for the whole float pool of IP addresses with whatever confidence is appropriate for all float pools of IP addresses with this number of class one samples, and the level of geographic commonality determined as follows. Therefore, if at least eighty-five percent (85%) of the class one samples agree at the state level of geographic hierarchy, that state will be established as the state for all of the IP addresses in the float pool of IP addresses may be considered at a calibrated level of confidence which may be higher or lower than eighty-five percent (85%).

As a more specific example, assume that for the one hundred IP addresses for which class one information is known, all of them have the same continent, ninety-five of them have the same country, ninety have the same state and region, eighty-eight have the same metropolitan area, forty have the same town, twenty have the same telephone exchange, and no two of them have the same ZIP code. Therefore, this float pool of IP addresses will be associated with the place record or lineage vector represented by the specific major metropolitan area to which the eighty-eight samples belonged, since this is the smallest level of the geographic hierarchy for which the single most common value represented at least eighty-five percent (85%), if eighty-five percent (85%) is used as the threshold, of the class one records sampled. Thus, any user having an IP address within this float pool of IP

addresses will be considered to have the place equal to that most common major metropolitan area, and with confidences at each other level of the geographic hierarchy as determined to be typical of all float pools of this type, specifically sharing (a) a similar number of class one sample records, and (b) the smallest level of the geographic hierarchy in common for eighty-five percent (85%) or more of the records being the major metropolitan area.

As another example, suppose a float pool of IP addresses generated using the boundary float pool technique 140 found twenty-six class one entries or samples in Illinois, U.S.A., one in Michigan, U.S.A., and one in Florida, U.S.A. Under the threshold system, each user of an IP address in the float pool of IP addresses would be considered to be in the state of Illinois, U.S.A., at a probability or confidence level appropriate to pools of this specific type and sub-type, based on number of samples and the level of the geographic hierarchy in common to within the outlier threshold (in this major metropolitan area). The actual state-level confidence level might be higher or lower than the outlier threshold of eighty-five percent (85%), depending on the result of the calibration process 175, but is typically slightly higher than eighty-five percent (85%), since outlier data often represents errors in data collection or data processing, or individuals who have moved or reported alternate addresses. If the outlier threshold were set at one hundred percent (100%), *i.e.*, the float pool of IP addresses is associated with the level in the geographic hierarchy which is in common to all class one records in the pool, each user of an IP address in this float pool of IP addresses would be considered to be in the United States of America, since the class one records varied in region, state, major metropolitan area, town, and ZIP code. The association of the pool of IP addresses with the whole of the USA is very accurate, but of considerably less utility since no other levels of the geographic hierarchy would be determined.

The threshold technique for generating confidence levels for technique lineage vectors for the float pool technique reduces the influence of "outliers" in the determination of confidence levels. Outliers are users who, for one reason or another, generate class one information that deviates from the majority of the other class one samples in a given float pool. For example, users may provide inaccurate or false name and address information class one information is collected. In addition, many users may be in metropolitan areas that span one or more states, countries, area codes, etc., such as the metropolitan area of New York City, U.S.A. Another reason for outliers is that a person may be visiting another city but using his/her own home address. Since the person is visiting another city, presumably they are using an IP address block which serves that city. Outlier handling provides a means for dealing with users that are located close to each other geographically, but who may occasionally provide address information in different areas at different levels of geographic hierarchy.

As previously described above, float puddles are generally given a lower confidence level for each level of geographic hierarchy than the float pool to which the float puddles are attached. For example, if a float pool is given an eighty-five percent (85%) confidence level at the state/province level of geographic hierarchy, float puddles adjacent the float pool might be given a confidence level of
5 seventy percent (70%) for the state/province level of geographic hierarchy. Empirical evidence and testing that a drop in confidence level of no more than eighteen percent (18%) is found for float puddles relative to adjacent float pools and is often much less. Therefore, a float puddle is preferably given the same geographic location as is the float pool to which the float puddle is adjacent, albeit at a lower confidence level for level of geographic hierarchy.

10 Fixed pools of IP addresses may be treated in a fashion similar to float pools of IP addresses. That is, if a fixed pool of IP addresses exists, all class one samples in the block of aligned IP addresses in the fixed pool are treated as if they were a float pool. The levels of geographic hierarchy that are at least eighty-five percent (85%) (other threshold levels may be used) invariant across the class one
15 samples are used to set the confidence levels for users of all of the IP addresses in the fixed pool. In this manner, the preferred technique of using such float puddles allows for the relatively successful geographic location determination of individuals using IP addresses in ranges for which no class one training information or data is available.

The place and reason (technique) lineage vectors may be used during the step 40 of the method
30 and will be discussed in more detail below.

20 During the step 38, a request is created for geographic location information for a particular user and, more specifically, the IP address associated with the user. The request may come from a variety of sources, including a banner server looking to send an advertisement or banner for display to a user.

25 Many selection methods are possible during the step 40. For example, the sub-technique or reason lineage may be chosen during the step 40 that has the highest confidence level at a specified level of geographic hierarchy. Alternatively, the sub-technique or reason lineage may be chosen during the step 40 that has the highest total of the confidence levels for all or some of the levels of geographic hierarchy. As a further alternative, the sub-technique or reason lineage preferably chosen is the sub-
30 lineage vector at which variances occur. For example, suppose that three different sub-techniques or reason lineages are used during the step 34 to determine the geographic location of a user associated with or assigned to a specific IP address. If the reason lineage vectors agree as to the continent, country, state/province of the user, even if at different confidence levels, but differ as to the metropolitan area, the sub-technique is chosen during the step 40 that provides the highest confidence

level for the metropolitan area, even if one or both of the other sub-techniques had a reason lineage vector with a higher confidence level for the continent, country, and/or state/province levels of geographic hierarchy.

The method 30 may also choose during the step 40 a technique or a sub-technique based on a function or combination of place lineage vectors and technique or reason lineage vectors. Preferably, sub-techniques are used since sub-techniques are broken down into different levels of geographic hierarchy, as previously discussed above. For example, suppose a request is received that desires to determine the geographic location of a specific user currently associated with a specific IP address. Assume also that the user has the following reason lineage vector associated with the sub-technique or reason lineage of the boundary float pool technique 140 directed to the major metropolitan level and a sample size of sixteen or more class one records within the float pool of IP addresses:

Continent selected will be accurate with a one-hundred percent (100%) confidence.

Country selected will be accurate with a one-hundred percent (100%) confidence.

State/province selected will be accurate with an eighty percent (80%) confidence.

Major metropolitan area selected will be accurate with a thirty percent (30%) confidence.

Town selected will be accurate with a zero percent (0%) confidence.

ZIP code selected will be accurate with a two percent (2%) confidence.

Telephone exchange selected will be accurate with a fifty percent (50%) confidence.

Such a reason lineage vector might be the result of ten IP addresses in a float pool having nine class one records that are identical at the state/province level of geographic hierarchy. Also assume that the place lineage vector for "Colorado" (*i.e.*, the level of geographic hierarchy at which the reason lineage vector is directed) is as follows:

Continent is North America with a one-hundred percent (100%) confidence.

Country selected is U.S.A. with a one-hundred percent (100%) confidence.

State/province is Colorado with a one-hundred percent (100%) confidence.

Major metropolitan area is Denver with a seventy percent (70%) confidence.

Town is Westminster with a zero percent (0%) confidence.

ZIP code is 80021 with a zero percent (0%) confidence.

Telephone exchange is 303 with an eighty percent (80%) confidence.

The place lineage vector and the reason lineage vector can be combined by multiplying each of their corresponding confidences, resulting in the following:

Continent is North America with a one-hundred percent (100%) confidence.

Country is U.S.A. with a one-hundred percent (100%) confidence.

State/province is Colorado with a eighty percent (80%) confidence.

Major metropolitan area is Denver with a twenty-one percent (21%) confidence.

Town is Westminster with a zero percent (0%) confidence.

ZIP code is 80021 with a zero percent (0%) confidence.

Telephone exchange is 303 with a forty percent (40%) confidence.

5 Other pairs of reason lineage vectors and place lineage vectors for a specified user can be multiplied in a similar fashion and the resulting vectors compared as similarly described above to provide a response during step 42.

10 Once the geographic location determination technique to be used to answer the request received during step 38 is determined during the step 40, the geographic location information can be returned to the requester during step 42 based on the information determined by the selected technique.

15 While the method 30 preferably creates or otherwise uses confidence levels or percentages for reason lineage and place lineage vectors, such percentages are largely based on empirical evidence. Therefore, it is desirable to have the ability to calibrate the method 30 to improve accuracy of the confidence levels for the method 30. In a general sense such a calibration for the method 30 could work as follows. First, class one records are gathered for users during step 32 and a targeting system is built using the method 30 as previously described above. Initial estimates for confidence levels or percentages can be used at each level of geographic hierarchy. Second, new class one records are gathered that differ from the class one records previously used. Such new class one records used for calibration are preferably distinct from the class one records used for the initial training of the system using the method 30. For the users associated with the new class one records, the method 30 is used to determine their location using all possible reason lineages. That is, assume that the new class one records do not exist and, using a known IP address of a class one individual for which one of the new class one records exists, try to predict where the users associated with the new class one records are geographically located. Third, the geographic location results of the second step are compared with the actual geographic location information gleaned from the new class one records. The results of the comparisons are used to adjust the confidence levels in the reason lineage vectors appropriately depending on how well the reason lineage vectors predicted the geographic locations of the users associated the new class one records. Fourth, the calibration process is repeated periodically or when otherwise desired as new sets of class one records are obtained that do not contain previously obtained class one records.

20

25

30

A more specific implementation of a calibration process 200 for use in determining accuracy of geographic location determination techniques is illustrated in Figure 15 and the process 200 can be used with or as part of the method 30. In the calibration process 200, a specific geographic location determination technique or sub-technique, both of which can be referred to as reason lineages, is

selected during the step 202. The technique selected could be, for example, the restricted range float pool technique 130 or the boundary float pool technique 140 previously discussed above. For the technique selected during the step 202, a particular instance or result of the technique is then identified and selected during the step 204. For example, if the boundary float pool technique 140 is selected

5 during the step 202, the pool of IP addresses illustrated by the histogram 180 in Figure 10 may be selected during the step 204 as a particular instance or result of the use of the boundary float pool technique. During step 206, a set of class one records associated with the pool of IP addresses selected during the step 204 is selected which could be targeted using the technique selected in the step 202. During step 208, without using any information in the class one records selected during the step 206,

10 the geographic locations of individuals with IP addresses corresponding to the class one records are determined using the technique or sub-technique selected during the step 202. Thus, during step 208, the geographic locations for individuals for whom class one record information is known are determined using only the IP addresses associated with the individuals and without using the class one record information for the individuals. During step 210, for each class one record selected during the step

15 206, and for each level of geographic hierarchy, the true location of the individual as determined from the individual's class one record is compared against the geographic location of the individual determined during the step 208 by using the technique selected during the step 202. During step 212, the percentage of comparisons or results which are correct are determined at each level of geographic hierarchy. After step 212, if more instances of the use of the technique chosen during the step 202 exist,

20 the process 200 preferably, but optionally, returns to step 204 to repeat the process 200 for the new instance of the use of the technique or sub-technique chosen during the step 202. After the steps 204, 206, 208, 210 are repeated for a desired number of instances of use of the technique selected during the step 202, during the step 214 the process 200 averages the results for each operation of the step 212 for the technique or sub-technique selected during the step 202, thereby providing an accuracy level of the

25 technique or sub-technique selected during the step 202 that can be used to establish confidence levels for the method 30. The process 200 can then be repeated as desired for other geographic location determination techniques or sub-techniques which could have been selected during the step 202.

The foregoing description is considered as illustrative only of the principles of the invention. Furthermore, since numerous modifications and changes will readily occur to those skilled in the art, it

30 is not desired to limit the invention to the exact construction and process shown and described above. Accordingly, all suitable modifications and equivalents may be resorted to falling within the scope of the invention as defined by the claims which follow. Furthermore, the words "comprise," "comprises," "comprising," "include," "including," and "includes" when used in this specification and in the following claims are intended to specify the presence of stated features, integers, components, or steps, but they

do not preclude the presence or addition of one or more other features, integers, components, steps, or groups thereof. In addition, the terms "computer" and "terminal" should not be interpreted in any limiting way and both include and encompass all types of client devices usable in a client/server environment or other devices, such as cable set top boxes, used in cable or other communication networks. The term "technique" as used in the following claims refers to, but is not limited to, each of the techniques, sub-techniques, and reason lineages previously described above.

It should also be noted that, while the techniques 98, 110, 130, 140, 160 previously discussed above relate cookies to IP addresses, the cookies associated with a particular IP address do not necessarily need to be completely or even partially identical. That is, so long as a mechanism or protocol is established such that IP addresses or cookies associated with other cookies or IP addresses, respectively, can be monitored, each of the techniques 98, 110, 130, 140, 160 can work properly and the use of cookies as described above does not imply that the cookies associated or used with a particular IP address or group of IP addresses are completely or even partially identical, that the cookies have a specific or particular structure or format, that the cookies contain specific or predefined information, or that the cookies or other distinct identifiers are used, related to, identified with, assigned to, sent or served to, or associated with a particular user or IP address or group of IP addresses in any predefined, set, or specific way or manner. Also, while cookies have been described throughout as usable as specific or distinct identifiers associated with IP addresses or other computer network addresses, specific or distinct identifiers could also be or include account numbers, equipment serial numbers, microprocessor identification numbers, cable set-top box addresses, etc. Furthermore, the method 30 and each of the techniques 98, 110, 130, 140, 160 can be used with other kinds of communication or cable networks and they are not limited to only computer networks or networks based on internet protocols. Thus, the method 30 and each of the techniques 98, 110, 130, 140, 160 can also be used with cable and other network addressing protocols or schemes.

Claims

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. A method for determining geographic location of a user connected to a network,
5 comprising:

collecting information regarding geographic location of a first set of one or more users of the network, each individual user in said first set having at least one network address that can be associated to said individual user's geographic location information;

determining a pool of network addresses, said pool of network addresses containing at
10 least one network address associated with a particular user in said first set of users; and

establishing at least a portion of said particular user's geographic location information as geographic location information for all users of network addresses in said pool of network addresses.

2. The method of claim 1, wherein said collected geographic information includes class
15 one information.

3. The method of claim 1, wherein said network addresses include internet protocol (IP) addresses.

4. The method of claim 1, including providing geographic location information for a user of a specific network address in said pool of network addresses to a requestor of geographic location
20 information associated with said specific network address.

5. The method of claim 4, wherein said geographic location information includes multiple levels of geographic hierarchy.

6. The method of claim 5, including establishing a confidence level for each level of geographic hierarchy for said user of said specific network address.

25 7. The method of claim 6, including calibrating said confidence levels.

8. The method of claim 7, wherein said calibrating said confidence levels includes collecting information regarding actual geographic location of a second set of one or more users of the network, each individual user in said second set having at least one network address that can be associated to said individual user's geographic location information.

30 9. The method of claim 8, wherein said calibrating said confidence levels includes determining geographic location of at least one selected user in said second set of users based on said selected user's network address.

10. The method of claim 9, wherein said calibrating said confidence levels includes comparing said selected users determined geographic location information with said selected user's actual geographic location information.

11. The method of claim 10, wherein said calibrating said confidence levels includes
5 establishing a percentage of confidence representative of said comparison of geographic location information.

12. The method of claim 6, including establishing a percentage of confidence of accuracy for each said level of geographic hierarchy.

13. The method claim 1, wherein said pool of network addresses includes contiguous or
10 consecutive network addresses.

14. The method of claim 1, wherein said geographic location information includes multiple levels of geographic hierarchy.

15. A method for determining a pool of network addresses, comprising:

collecting geographic location information for a user associated with at least one
15 network address;

determining a first network address for which a cookie or other specific identifier is or has been associated and a second network address, distinct from and higher than said first network address for which said cookie or other specific identifier also is or has been associated;

associating a counter to at least one network address between said first network
20 address and said second network address; and

incrementing each of said at least one counters.

16. The method of claim 15, wherein each of said at least counters is incremented in an equal amount.

17. The method of claim 15, wherein said collecting, determining, associating, and
25 incrementing steps are repeated for a plurality of distinct cookies or other specific identifiers.

18. The method of claim 17, wherein said incrementing said at least one counters creates a histogram.

19. The method of claim 15, wherein said network address associated with said geographic location information for said user is in a range of network addresses between and including said first
30 network address and said second network address.

20. The method of claim 15, wherein said collected geographic location information includes class one information.

21. The method of claim 15, including initializing each of said at least one counters to an initial value.

22. The method of claim 17, including initializing each of said at least one counters to an initial value.

23. The method of claim 22, including resetting at least one of said at least one counters to said initial value if said at least one counter is more than a first threshold number less than values for each counter associated with a network address within a range of network addresses of the network address associated with said one of said at least one counters.

24. The method of claim 23, wherein said first threshold number is three.

25. The method of claim 23, wherein said range of network addresses includes network addresses within a second threshold number of network address higher and lower than the network address associated with said one of said at least one counters.

26. The method of claim 15, wherein each of said network addresses are or include internet protocol addresses.

27. A method for determining a pool of network addresses, comprising:

selecting a range of network addresses;

selecting a first network address within said range of network addresses;

determining if a specific identifier has been associated with at least one network address in said range of network addresses that is higher than said first network address and if said specific identifier has been associated with at least one network address in said range of network addresses that is lower than said first network address; and

repeating said process for a second network address within said range of network addresses.

28. The method of claim 27, wherein said range of network addresses includes all possible network addresses.

29. The method of claim 27, wherein said range of network addresses includes only a subset of all possible network addresses.

30. The method of claim 27, including collecting geographic location information for a user associated with at least one network address in said range of network addresses.

31. The method of claim 27, including collecting geographic location information for a user associated with said first network address.

32. The method of claim 27, wherein said specific identifier is a cookie.

33. The method of claim 27, wherein said specific identifier forms or includes at least a portion of a cookie.

34. The method of claim 27, including determining, for each of said network addresses in said range of network addresses, whether there exists one or more specific identifiers that are associated

with a higher network address in said range of network addresses and a lower network address in said range of network addresses.

35. The method of claim 34, wherein each of said one or more specific identifiers is a cookie.

5 36. The method of claim 34, wherein each of said one or more specific identifiers forms or includes at least a portion of a cookie.

37. The method of claim 34, including collecting geographic location information for a user associated with an selected network address for which a specific identifier exists that is associated with network addresses in said range of network addresses both higher and lower than said selected
10 network address.

38. The method of claim 37, wherein, for each network address between said higher and lower network addresses, there exists at least one specific identifier that is associated with network addresses in said range of network addresses both higher and lower than said each network address between said higher and lower network addresses.

15 39. The method of claim 27, wherein said range of network addresses includes contiguous or consecutive network addresses.

40. The method of claim 27, wherein said range of network addresses includes non-contiguous or non-consecutive network addresses.

20 41. The method of claim 27, wherein each network address in said range of network addresses is or includes an internet protocol address.

42. A method for determining geographic location of a specific user connected to a connected to a network; comprising:

determining with a first technique a first possible geographic location of the specific user;

25 determining with a second technique a second possible geographic location of the specific user;

determining whether said first technique more accurately determined the specific user's actual geographic location or whether said second technique more accurately determined the specific user's actual geographic location; and

30 calibrating said first technique and said second technique.

43. The method of claim 42, wherein said first technique includes determining a first pool of internet protocol addresses, said first pool of internet protocol addresses including at least one internet protocol address associated with the specific user.

44. The method of claim 43, wherein said second includes determining a second pool of internet protocol addresses, said second pool of internet protocol addresses including at least one internet protocol address associated with the specific user.

45. The method of claim 43, wherein said first technique includes associating geographic information for an internet protocol address in said first pool of internet protocol addresses with all internet protocol addresses in said first pool of internet protocol addresses.

46. The method of claim 42, wherein said first technique includes a boundary float pool determination approach.

47. The method of claim 42, wherein said first technique includes a float pool determination approach.

48. The method of claim 47, wherein said first technique includes a range float pool determination approach.

49. The method of claim 47, wherein said first technique includes a restricted float pool determination approach.

50. The method of claim 42, wherein said first technique includes a proxy server float pool determination approach.

51. The method of claim 42, including receiving a request for geographic location information for the specific user.

52. The method of claim 51, including providing said geographic location information for the specific user to said requestor.

53. The method of claim 42, wherein said first technique includes gathering geographic location information for at least one user who has been connected to the network.

54. The method of claim 53, wherein said first technique includes determining a first set of internet protocol addresses that have been used or associated with said at least one user.

55. The method of claim 54, wherein said first technique includes determining a first set of cookies or other specific identifiers that have been used or associated with any internet protocol addresses in said first set of protocol addresses.

56. The method of claim 55, wherein said first technique includes determining a second set of internet protocol addresses that have been used or associated with any of said cookies or other specific identifiers in said first set of cookies or other specific identifiers.

57. The method of claim 56, wherein said first technique includes determining a second set of cookies or other specific identifiers that have been used or associated with any of said internet protocol addresses in said second set of internet protocol addresses.

58. The method of claim 57, wherein said first technique includes repeatedly determining sets of cookies or other specific identifiers and internet protocol addresses until such time as said sets of internet protocol addresses stabilizes for consecutive determinations.

59. A method of determining a pool of network addresses, comprising:

5 gathering geographic location information for at least one user who has been connected to the network;

determining a first set of network addresses that have been used or associated with said at least one user;

10 determining a first set of cookies or other specific identifiers that have been used or associated with any network addresses in said first set of network addresses;

determining a second set of network addresses that have been used or associated with any of said cookies or other specific identifiers in said first set of cookies or other specific identifiers;

15 determining a second set of cookies or other specific identifiers that have been used or associated with any of said network addresses in said second set of network addresses; and

repeatedly determining sets of cookies or other specific identifiers and network addresses until such time as said sets of network addresses stabilizes for consecutive determinations.

20 60. The method of claim 59, wherein cookies or other specific identifiers that have been seen or associated with a first network address and a second network address are removed from said sets of cookies or other specific identifiers if said first protocol address and said second protocol address differ by at least a threshold value.

61. The method of claim 60, wherein said threshold value is equal to 4,096.

25 62. The method of claim 59, including using said collected geographic targeting information to establish geographic location of a user using a network address that falls within said stable set of network addresses.

63. The method of claim 59, wherein each of said network addresses is or includes an internet protocol address.

64. A method of determining a pool of network addresses, comprising:

30 gathering geographic location information for at least one user who has been connected to the network;

determining a lower network address and a higher network address that have been used or associated with said at least one user;

determining a first set of one or more specific identifiers that have been used or associated with any network addresses in a range between said lower network addresses and higher network address;

5 determining a new lower network address and a new higher network address that have been used or associated with any of said one or more specific identifiers in said first set of one or more specific identifiers;

determining a second set of one or more specific identifiers that have been used or associated with any network addresses in a range between said new lower network address and set new higher network address; and

10 repeatedly determining sets of one or more specific identifiers and lower and higher network addresses until such time as said lower network address stabilizes and said higher network address stabilizes.

65. The method of claim 64, including using said collected geographic targeting information to establish geographic location of a user using a network address that falls in a range between and including said stable lower network address and said stable higher network address.

66. The method of claim 64, wherein each of said one or more specific identifiers is a cookie.

67. The method of claim 64, wherein each of said one or more specific identifiers forms or includes at least a portion of a cookie.

20 68. The method of claim 64, wherein each of said network addresses is or includes an internet protocol address.

1/14

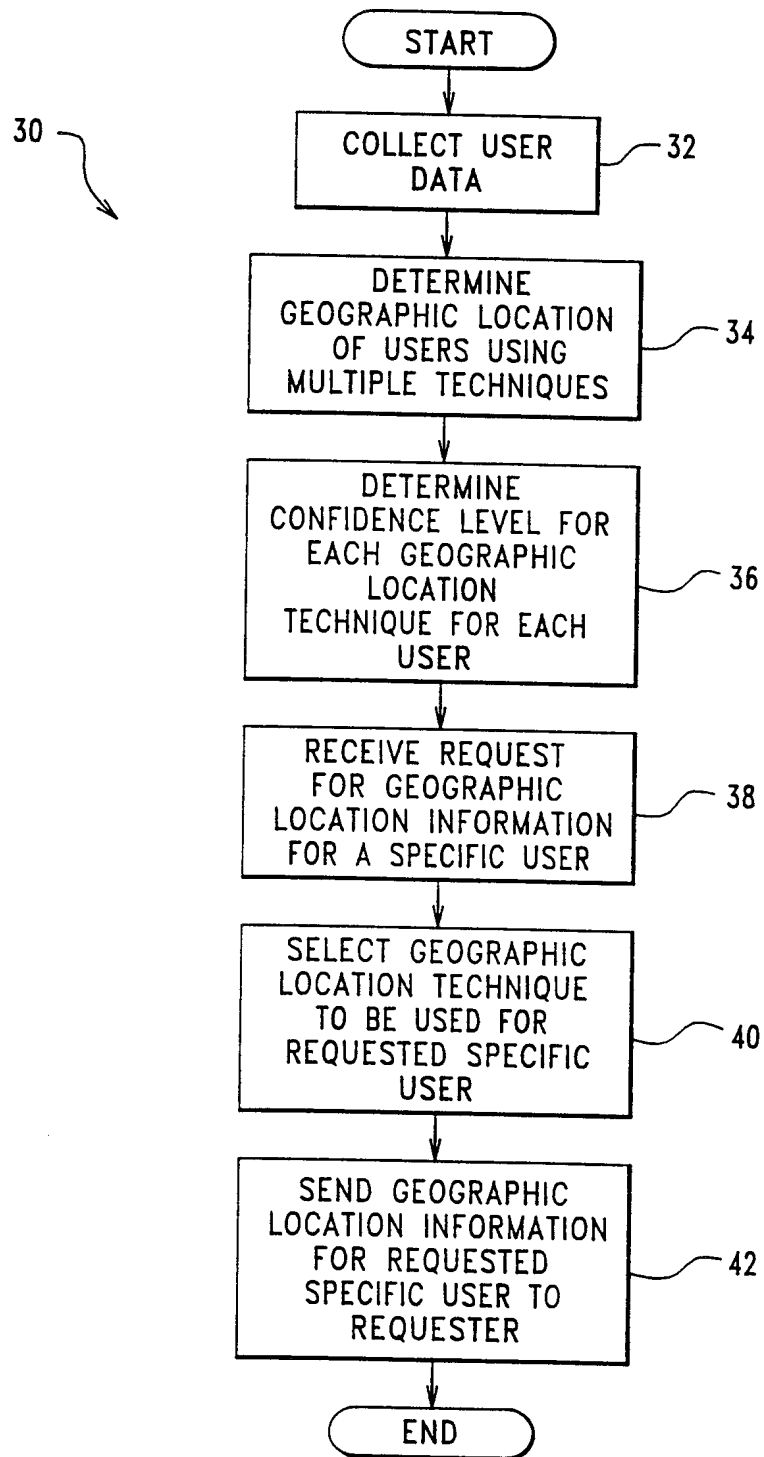


FIG. 1

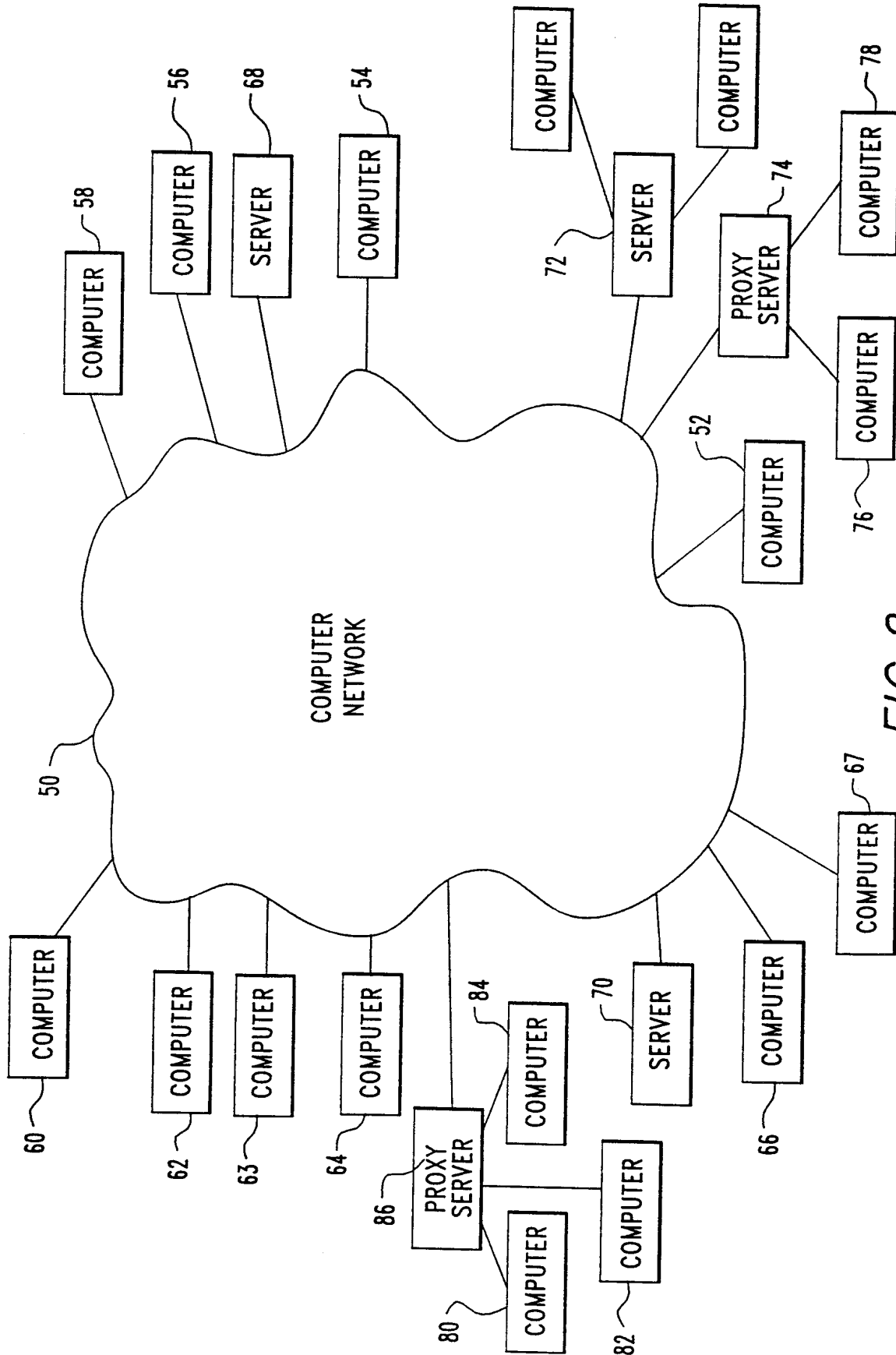


FIG. 2

3/14

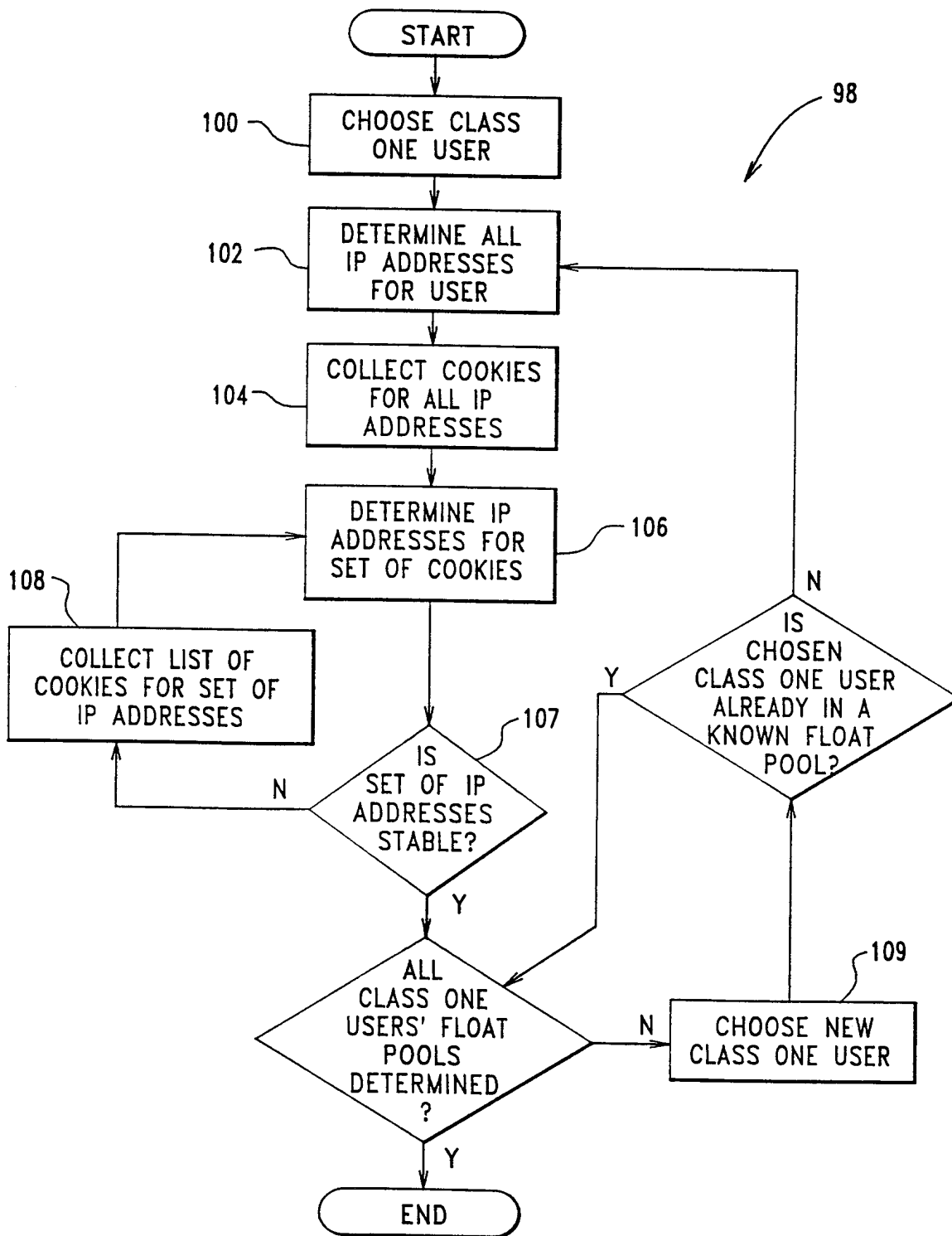


FIG. 3

4/14

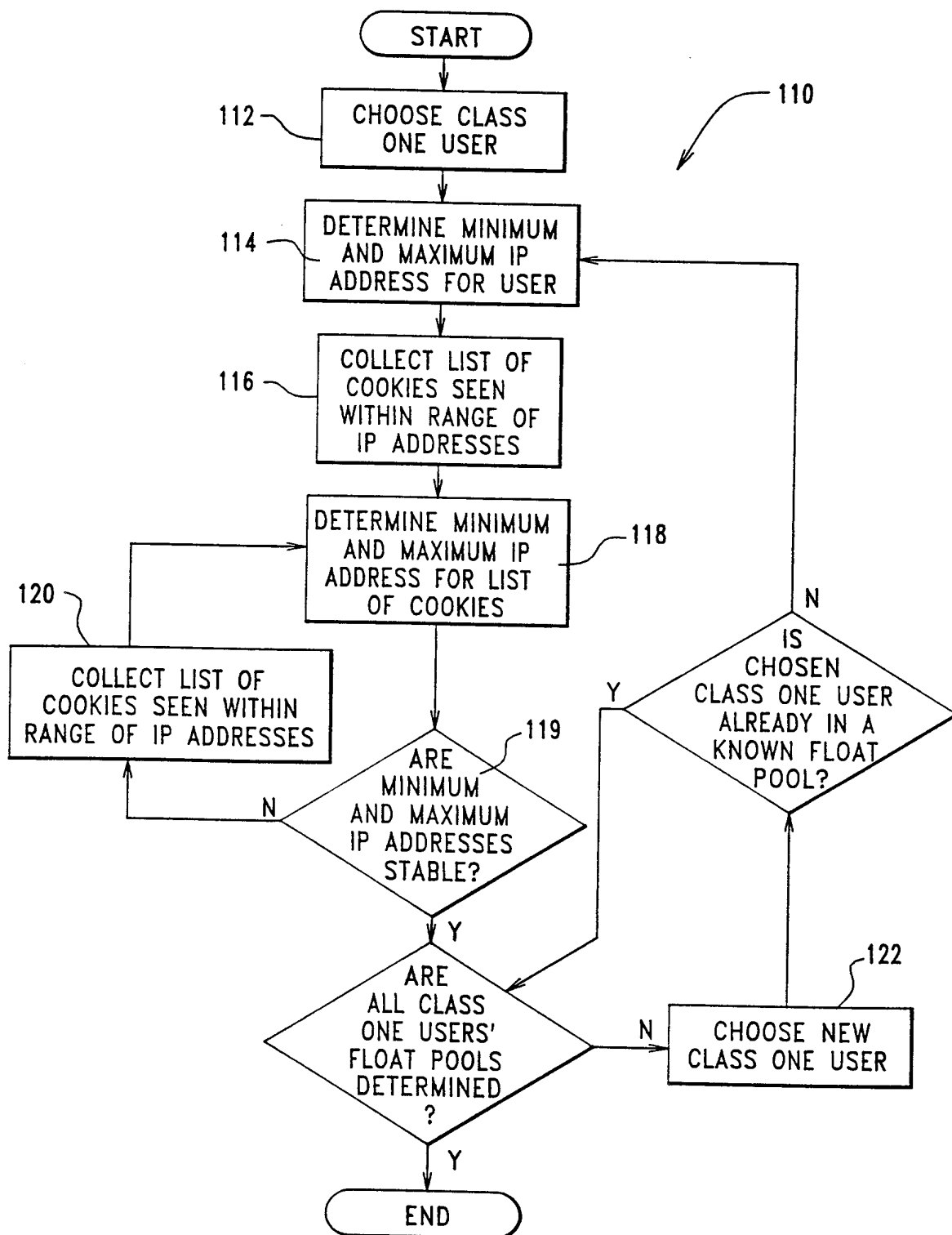


FIG. 4

5/14

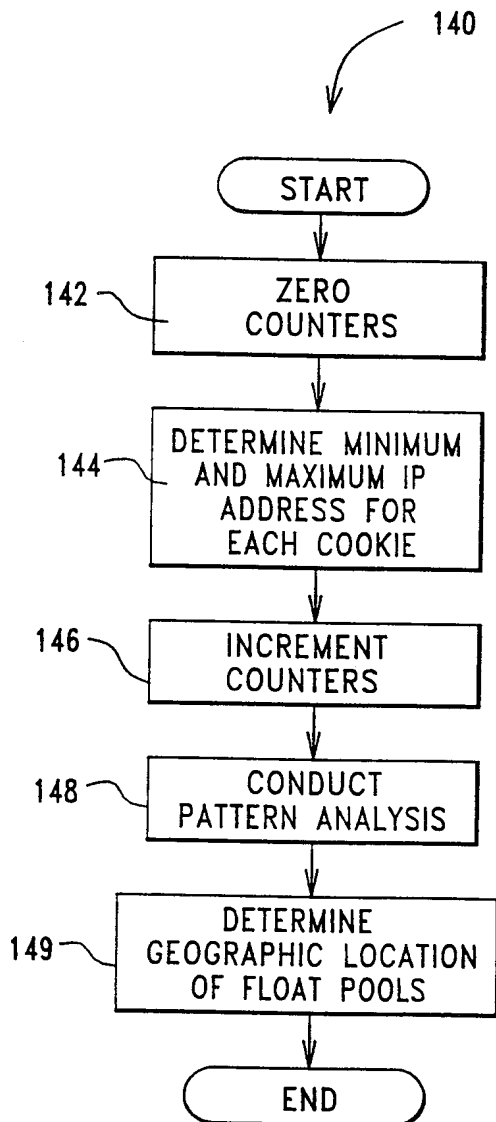


FIG. 5

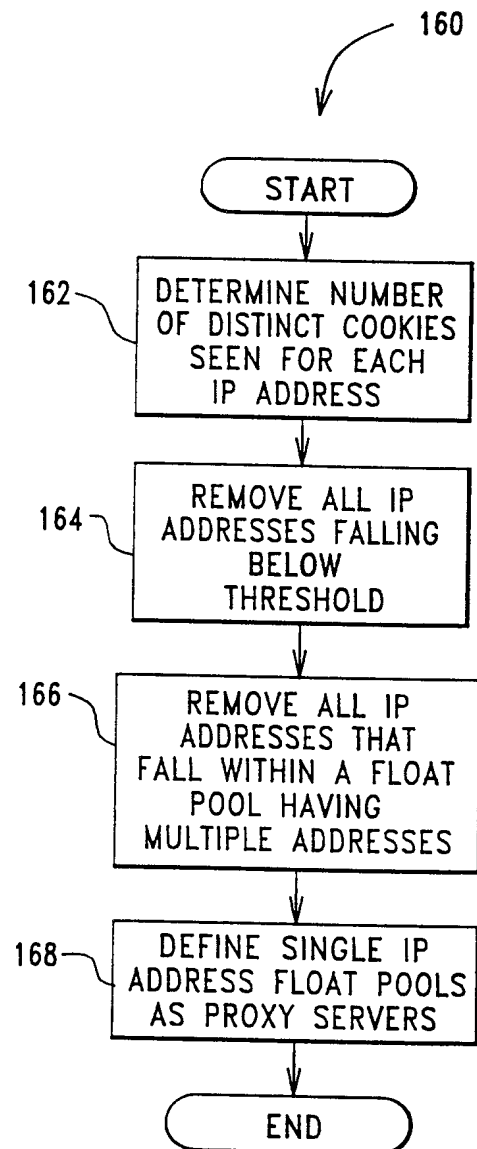


FIG. 9

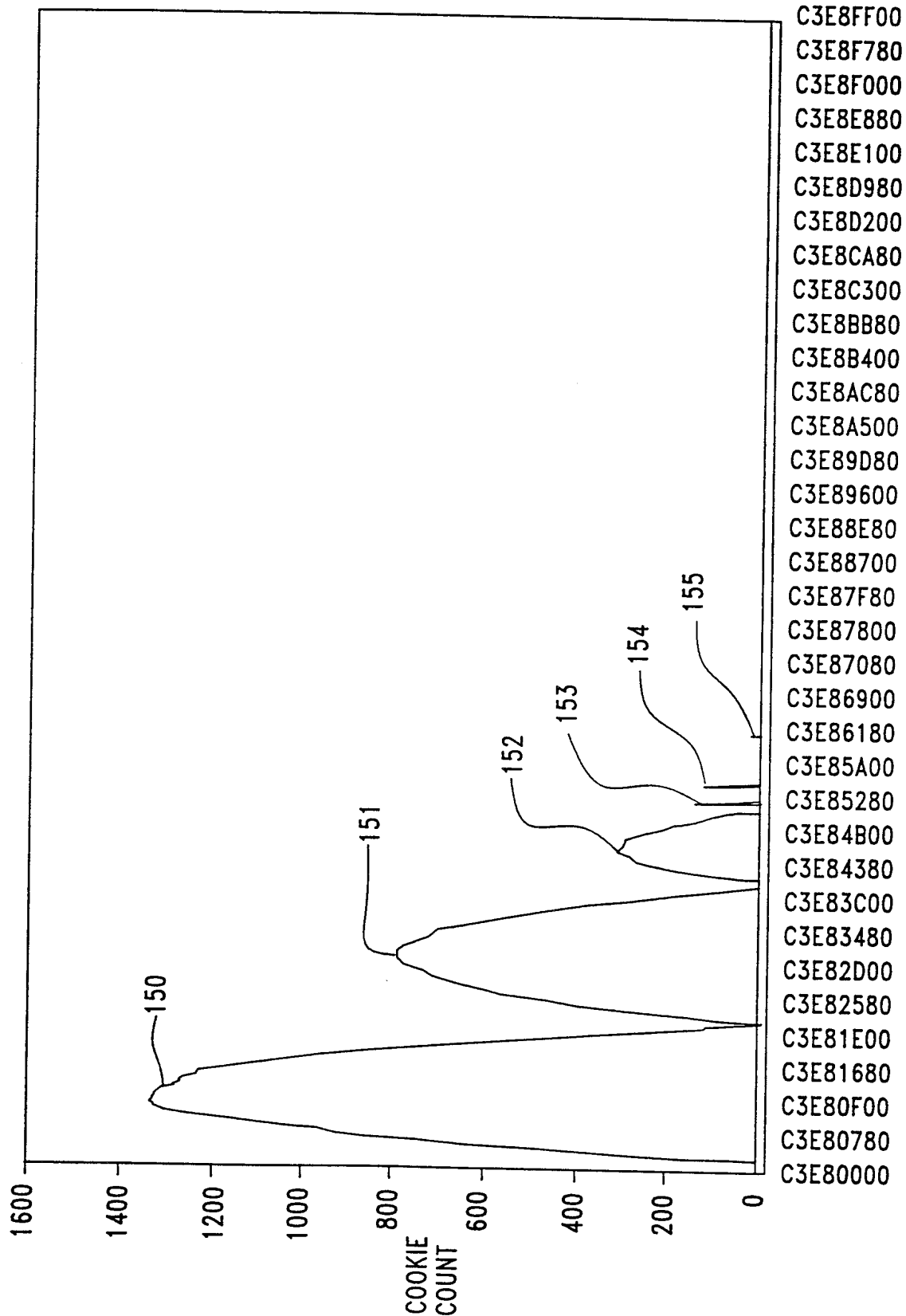


FIG. 6

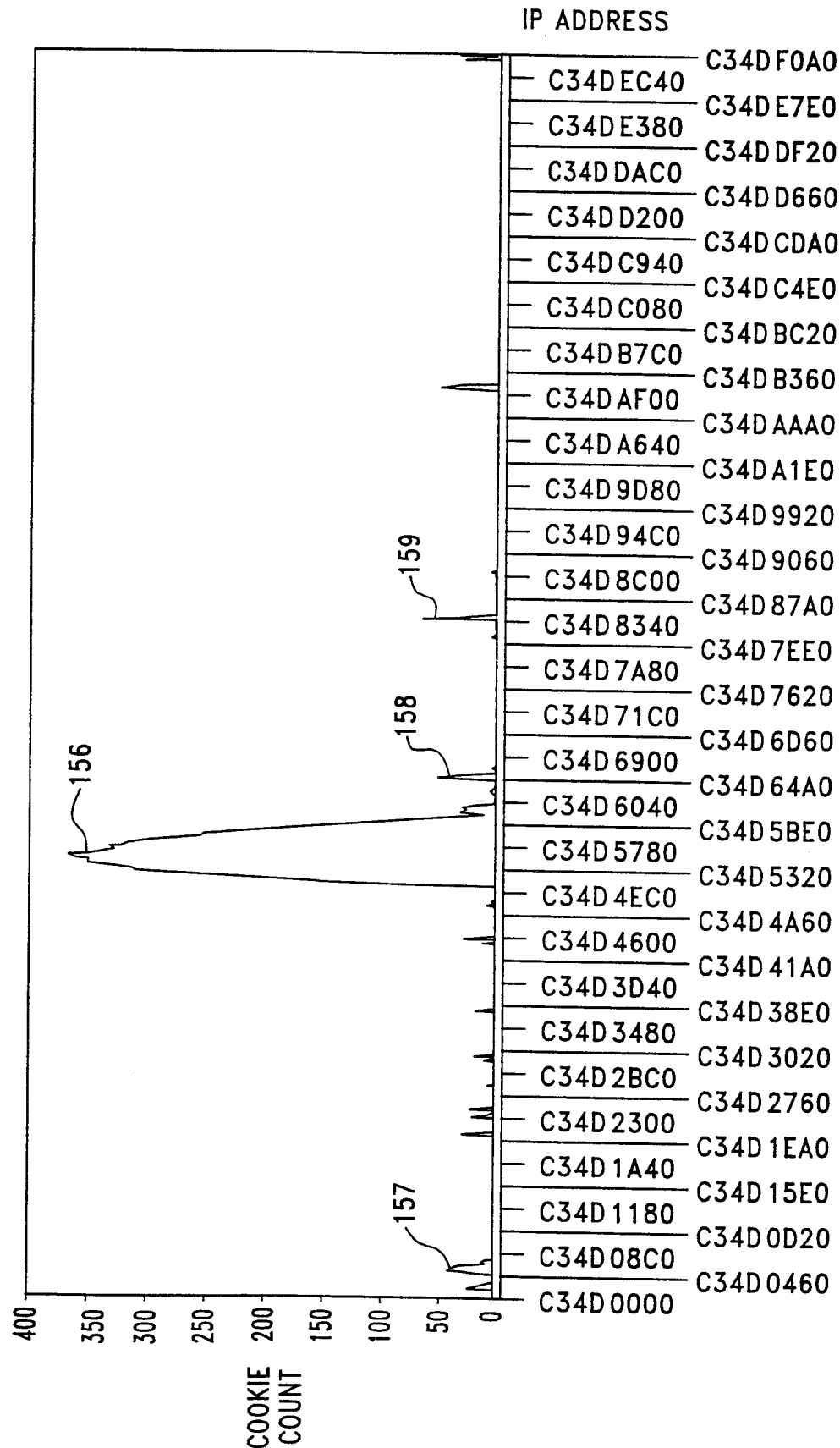


FIG. 7

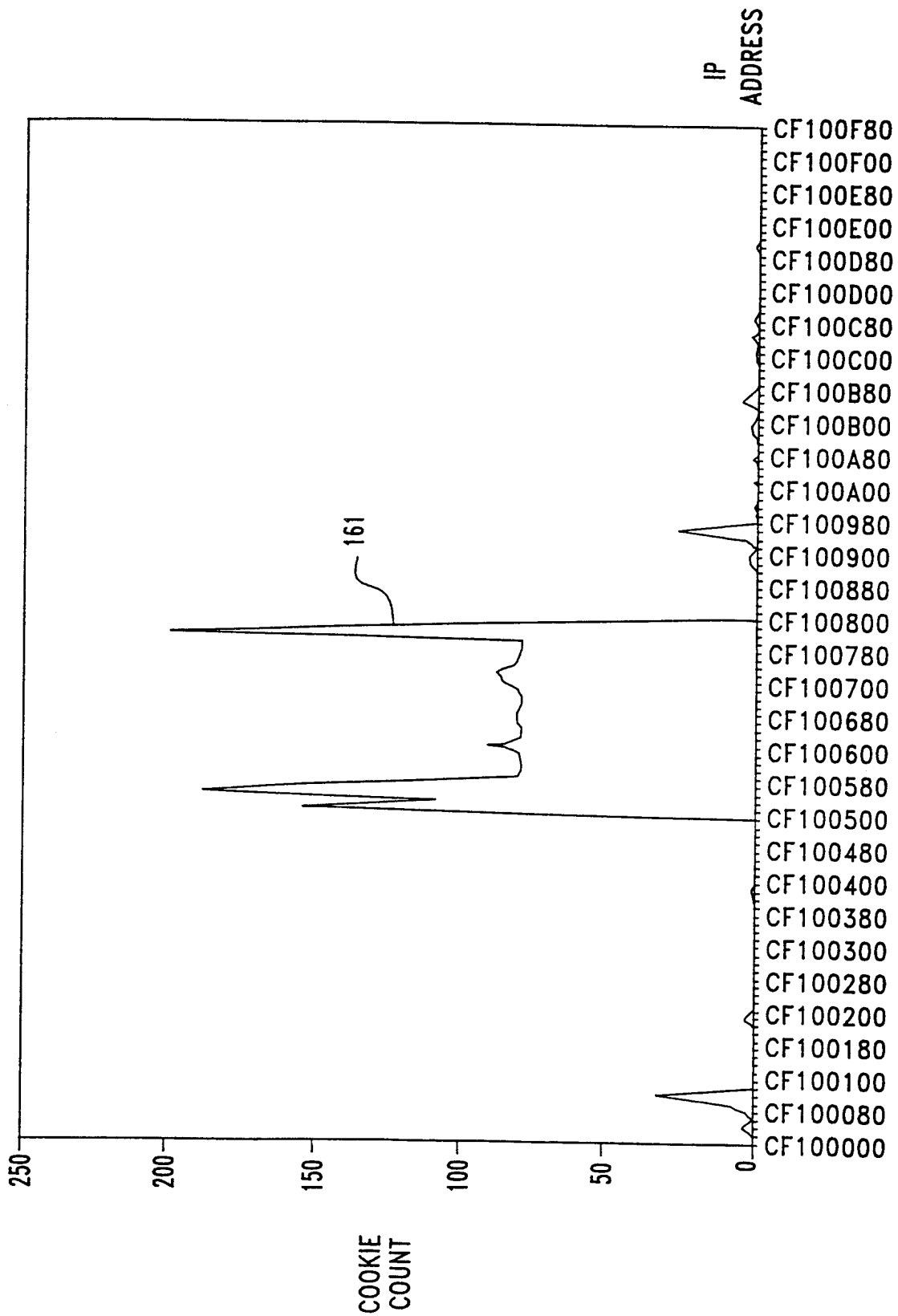


FIG. 8

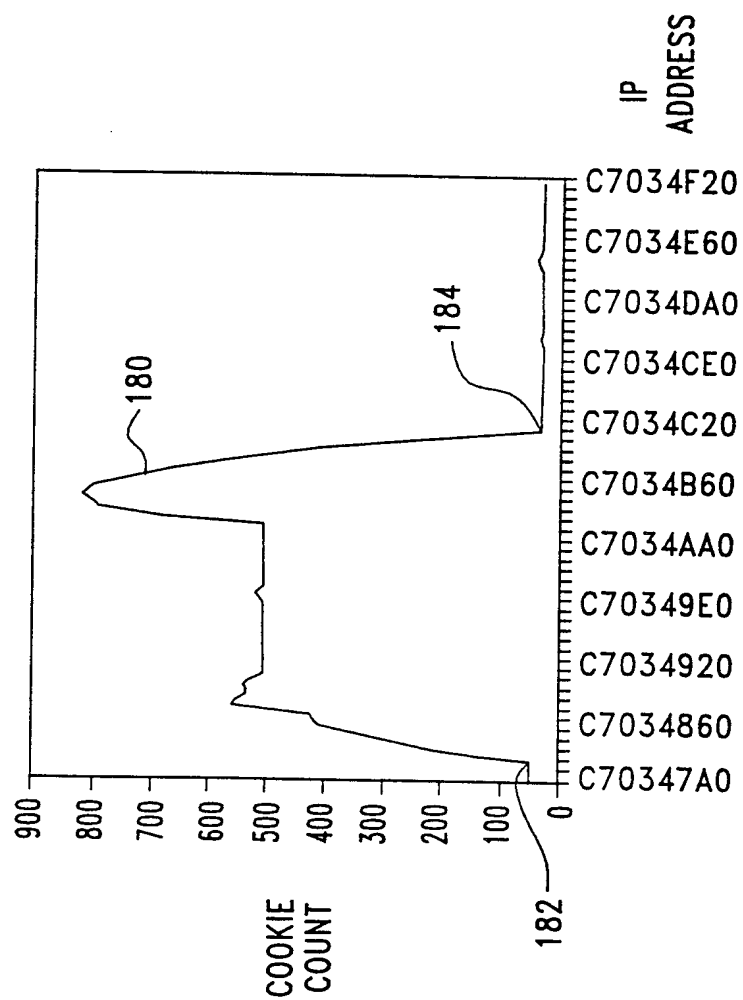


FIG. 10

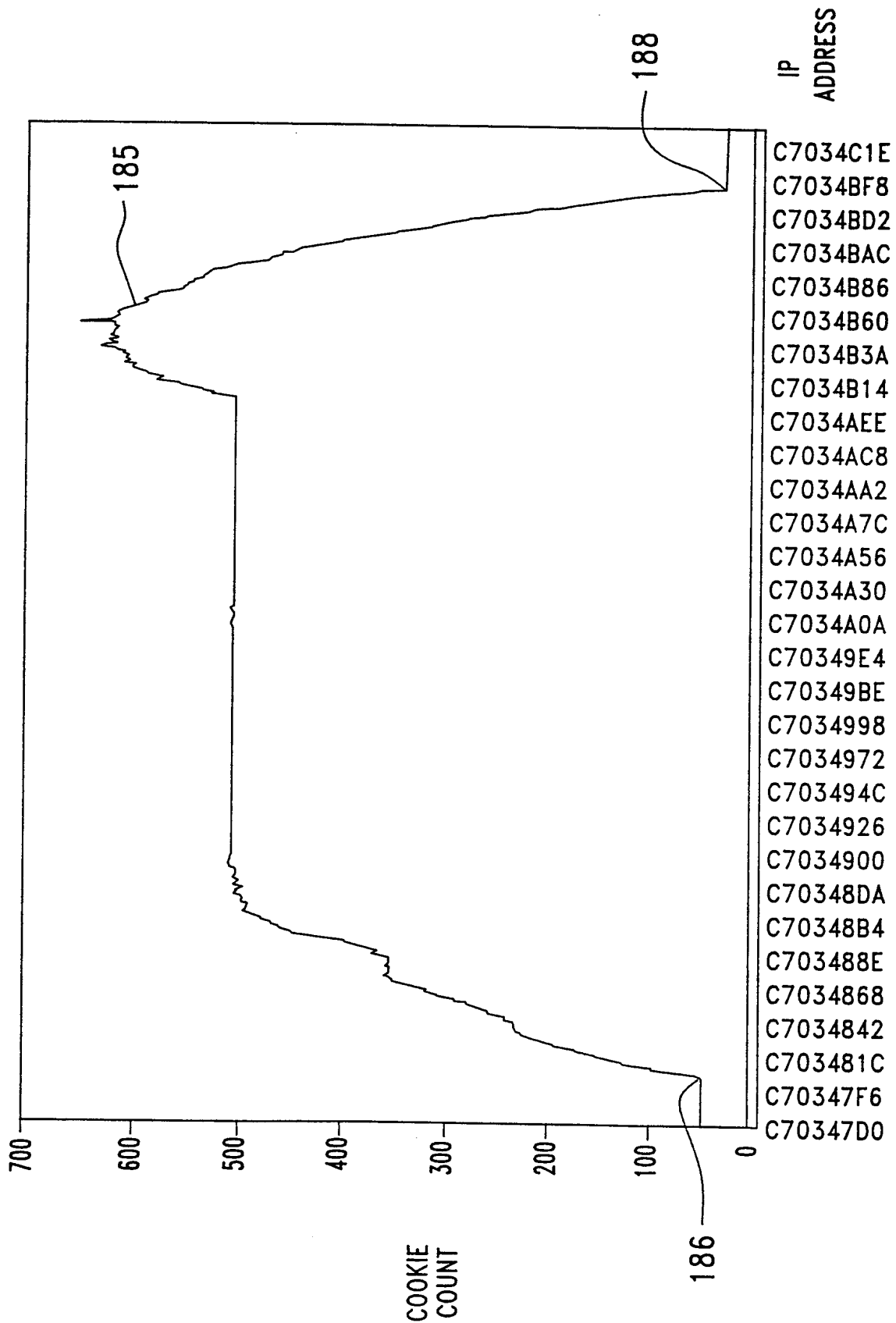


FIG. 11

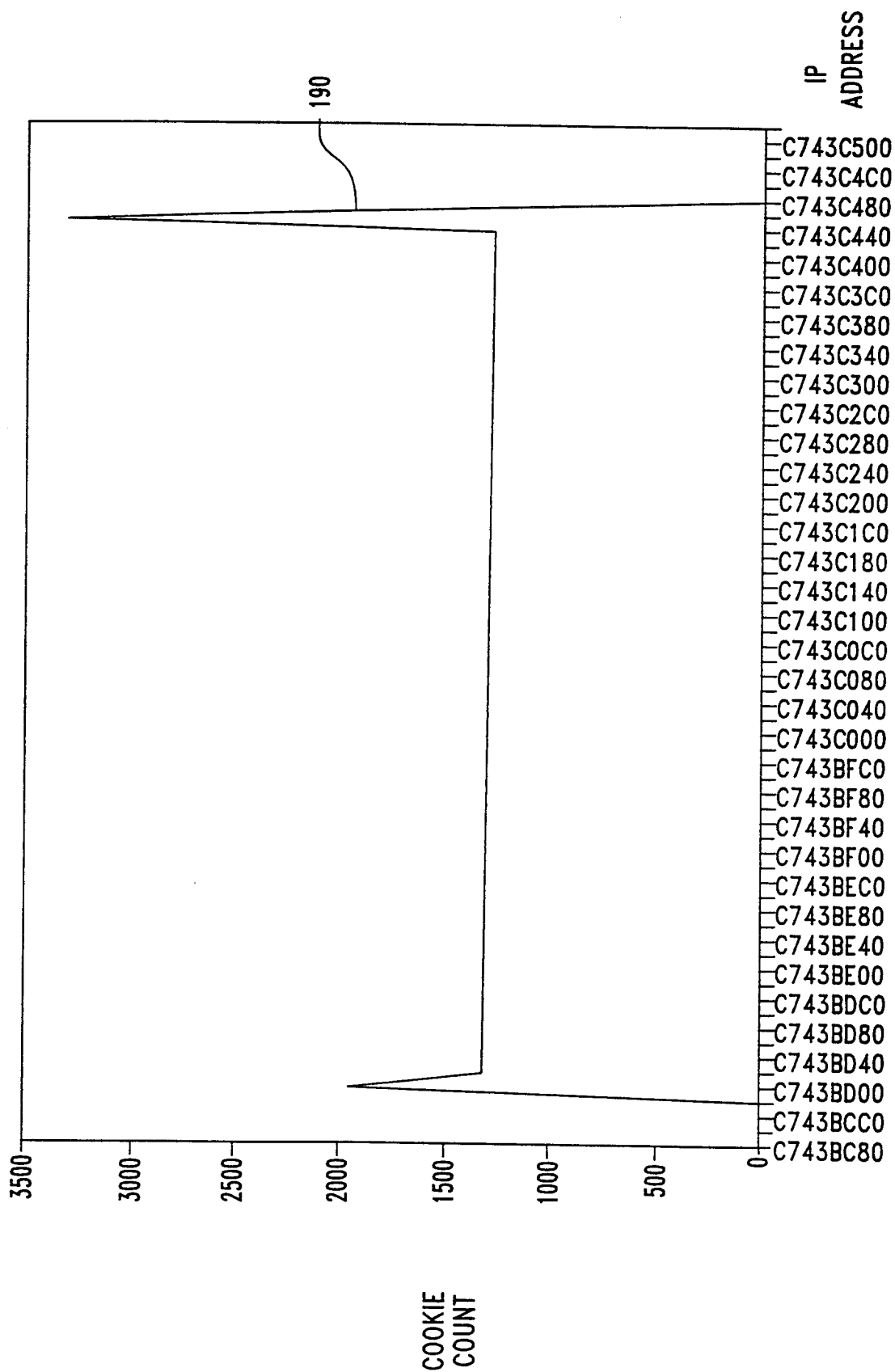


FIG. 12

12/14

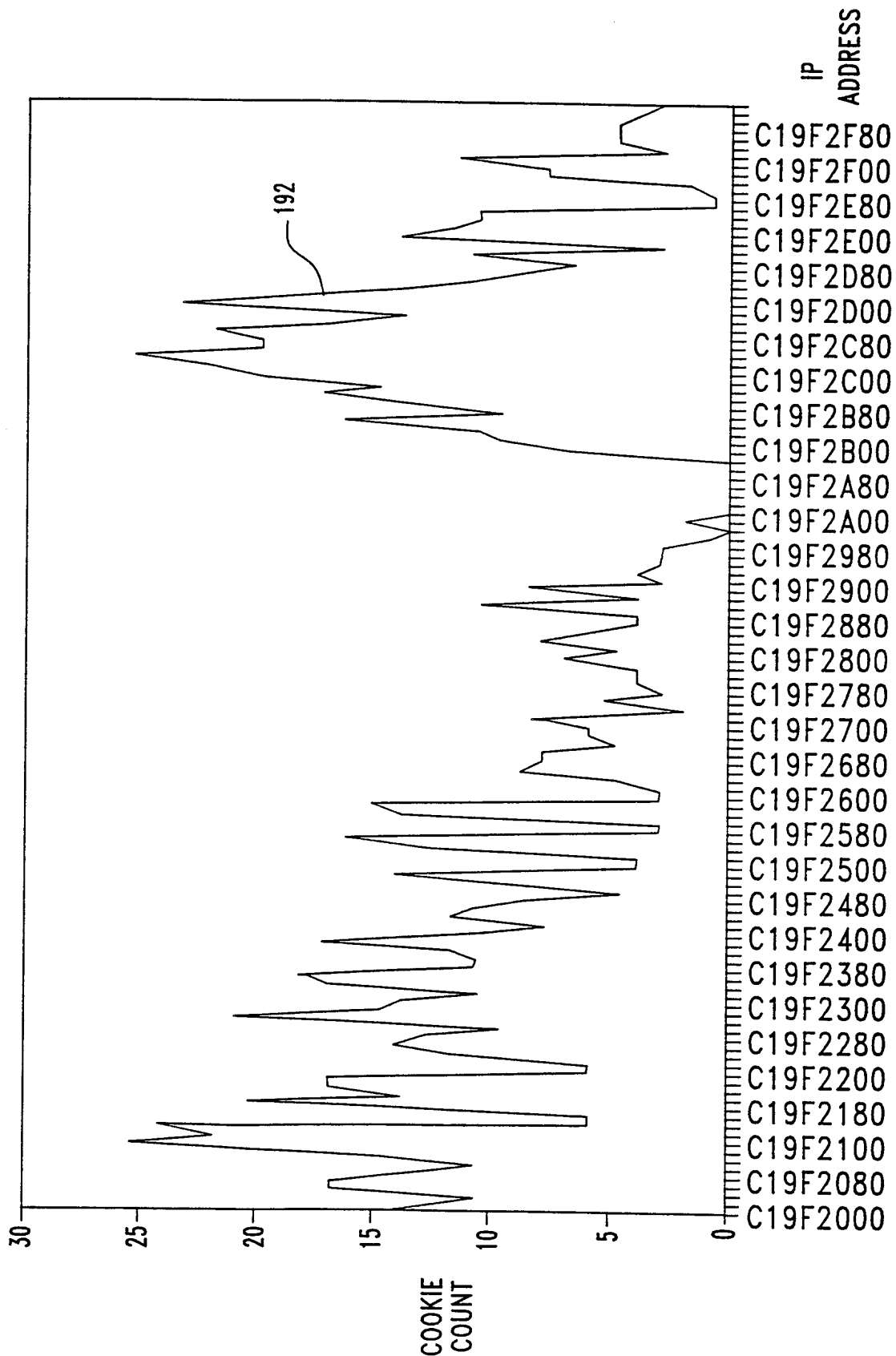


FIG. 13

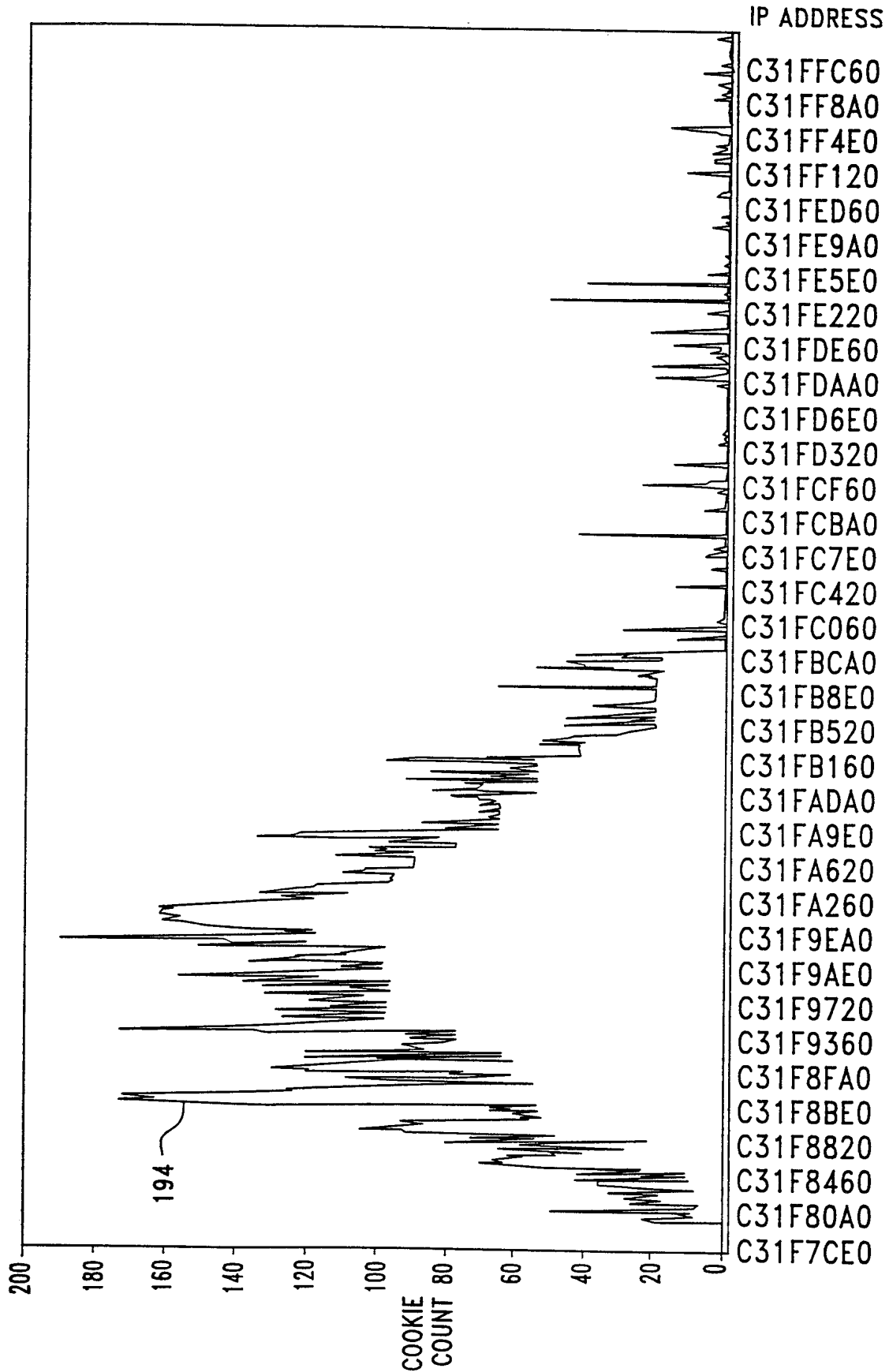


FIG. 14

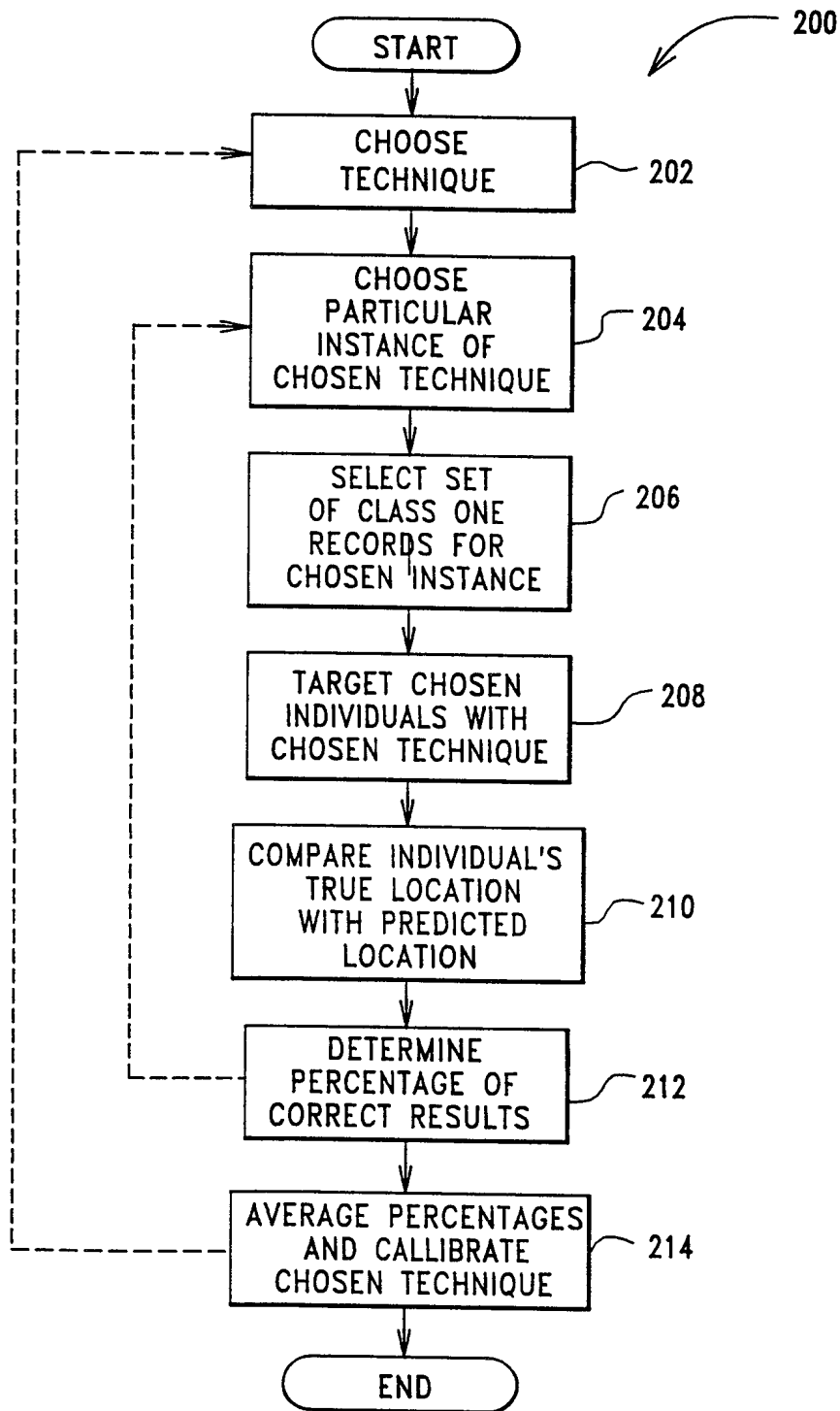


FIG. 15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/04934

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :GO6F 15/00, 15/16, 17/30, 151/00

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 345/327; 705/14; 707/10, 100, 102, 103, 501; 709/218, 219, 224, 238, 239

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,948,061 A (MERRIMAN ET AL.) 07 SEPTEMBER 1999, COL.1-8	1-68
Y, E	US 6,044,376 A (KURTZMAN, II) 28 MARCH 2000, COL. 1-5	1-68
Y, E	US 6,098,106 A (PHILYAW ET AL.) 01 AUGUST 2000, COL.1-14.	1-68

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	* & * document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 28 AUGUST 2000	Date of mailing of the international search report 22 SEP 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer GLEN BURGESS <i>James R. Matthews</i> Telephone No. (703) 305-4792

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04934

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

345/327; 705/14; 707/10, 100, 102, 103, 501; 709/218, 219, 224, 238, 239